PK/SH 8001

# REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188



| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE | 3. REPORT TYPE AND DATES COVERED |
|---|---|---|
| | 14 October 1994 | Final Tech. Report, 1 July 1991-30 June 199 |

**4. TITLE AND SUBTITLE**
Inverse and Control Problems in Electromagnetics
Final Technical Report

**5. FUNDING NUMBERS**
AFOSR Grant
91-0277

**6. AUTHOR(S)**
T. S. Angell and R. E. Kleinman

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Department of Mathematical Sciences
University of Delaware
Newark, DE 19716

**8. PERFORMING ORGANIZATION REPORT NUMBER**
AFOSR-TR- 95-0199

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
AFOSR/NM
Bldg. 410
Bolling AFB, DC 20332-6448



**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**
Approved for public release; distribution unlimited

**12b. DISTRIBUTION CODE**



**13. ABSTRACT** (Maximum 200 words)

This report summarizes work carried out in a number of specific areas of investigation in inverse scattering and optimal control problems in electromagnetics. The progress is briefly described and detailed results are included in an appendix. The major accomplishments include: the application of multi-criteria optimization techniques to problems in antenna design; the development of inverse scattering algorithms which use scattered field data in the frequency domain to reconstruct the shape, location and constitutive parameters of a scattering object; establishing the well-posedness of electromagnetic scattering problems with resistive or conductive boundary conditions; and derivation of new boundary integral equations for electromagnetic scattering from local distributions of a plane screen. In addition some new results on low frequency scattering have been found which establish the exact nature of the asymptotic expansion in two dimensions.

| 14. SUBJECT TERMS | 15. NUMBER OF PAGES |
|---|---|
| Antenna design, multicriteria optimization, inverse scattering, low frequency scattering, integral equations | 236 |
| | 16. PRICE CODE |

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| unclassified | unclassified | unclassified | |

FINAL TECHNICAL REPORT

TO

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

ON

AFOSR GRANT NO. 91–0277

# INVERSE AND CONTROL PROBLEMS IN ELECTROMAGNETICS


Name of Institution: Department of Mathematical Sciences
University of Delaware
Newark, DE 19716

Time Period: July 1, 1991 – June 30, 1994

Principal Investigators: Ralph E. Kleinman
Thomas S. Angell

Program Officer: Marc Jacobs

Final Technical Report

AFOSR Grant No. 91–0277

1 July 1991 — 30 June 1994

## 1. Research Objectives

The research conducted under this grant is a continuation of a long–term research effort devoted to the study of various aspects of direct and inverse electromagnetic scattering previously supported by AFOSR. The general goal of the program continues to be the establishment of a firm mathematical foundation and the development of algorithms based on such a foundation in which boundary and domain parameters are either to be recovered from scattering or radiation data or used as controls to optimize various functionals of the scattered or radiated fields. Such parameters include the shape of the boundary itself, functions defined on the boundary such as impedance, generalized impedance and generalized resistivity, as well as domain parameters such as conductivity and refractive index. The program, as described in the original proposal, is focussed on three specific areas of investigation: multicriteria optimization, generalized impedance boundary conditions, and inverse scattering techniques.

We summarize the nature of each of the particular problem areas and report on work done and in progress in each of the next four sections. In addition to this work, we have begun the writing of a monograph devoted to optimization methods in antenna theory which will be devoted, to a large extent, to the systematic exposition of the theory and computational results obtained with the support of several AFOSR grants. This monograph is being written in collaboration with Professor A. Kirsch

of the Universität Erlangen–Nürnberg in Germany.

## 2. Research Accomplishments and Current Status

### 2.1. Multicriteria Optimization

Many problems of applied interest in both the optimization of radiated fields and the identification of targets may be viewed as involving several performance criteria, any one of which may be taken as the primary cost functional which is to be optimized. A variety of such performance is evidenced in antenna problems as described in [3.2][1]. Other desirable characteristics, represented as functionals, are most often treated as constraints to be satisfied by an optimal solution, and some multiplier technique is used to produce an unconstrained problem.

However, the designation of one primary cost functional and the relegation of others to the status of constraints, is somewhat arbitrary. Indeed, a more direct approach is to consider such problems as multicriteria problems of optimization. To our knowledge, our use of multicriteria techniques is new in the fields of inverse scattering and control in electromagnetics. The ideas were first presented to the electromagnetic community at the Boulder URSI Meeting in 1992 [5.5].

We have prepared two manuscripts on this subject each of which includes both theoretical analysis as well as computational results. The actual computation of the manifold of Pareto optimal points gives the design engineer a range of choices making the trade–offs between different optimal choices explicit. The first paper describing these results, [3.5], *Multicriteria Optimization in Antenna Design* ap-

[1] Numbers in parenthesis refer to papers and presentations listed in sections 3 and 5.

peared in 1992. A second paper [5.12] *Multicriteria Optimization in Arrays* was presented at the JINA 92 meeting in Nice, France in November, 1992. This paper addresses the use of such methods for antenna arrays and compares these results with the well-known Dolph-Tchebyscheff result. We presented these results at a seminar at Rome Laboratories, Hanscom AFB in January, 1993. At that time, it became clear that an array problem previously considered by R.A. Shore of Rome Laboratories could also be treated by multicriteria methods. It was agreed that we would collaborate on the application of the multicriteria approach to this problem. This work is ongoing.

The work described above includes numerical computations for problems involving both arrays and conformal antennas. Related to these problems is that of maximizing the power in a preassigned sector of the far field. We considered this problem several years ago. In the present grant period, we returned to that problem and, in collaboration with B. Vainberg of the University of North Carolina at Charlottesville, we have been able to use asymptotic methods to *characterize* the optimal surface current in terms of a graph norm for the Neumann-to-Dirichlet operator. In doing so, we get an *explicit* representation for this operator. These results were presented [5.20] and will appear in [3.12].

## 2.2. Generalized Boundary Conditions

Under the present grant, we initiated a study of the well-posedness of resistive and conductive boundary value problems for the acoustic case. Under the present grant, we have completed a paper *The Conductive Boundary Condition for Maxwell's Equations* in collaboration with A. Kirsch [3.4]. The results were reported

at the IEEE/APS URSI International Symposium in Chicago in July, 1992, [5.8].

These conditions, intended to model thin layer behavior, are neither pure boundary conditions nor full transmission conditions, and involve using variable resistivity or conductivity to model such layers. An alternative is the use of higher order or generalized impedance conditions i.e., boundary conditions which involve differential operators of higher than the order of the differential equation.

In collaboration with S. Przeździecki of the Polish Academy of Sciences, we attempted a rigorous derivation of such conditions for the electromagnetic scattering in the case of a plane stratified medium. Using Fourier transform techniques the problem was transformed into one involving a set of transmission line equations. A preliminary version of this work was presented in [5.6] but some details of this work remain to be clarified before a manuscript can be completed.

### 2.3. Inverse Problems

We have pursued three lines of research on this topic. The first is the development of an efficient computer algorithm for a variant of the shape identification method based on complete families of solutions which we developed under the previous grant. Work with J. Jiang, a postdoctoral fellow, has yielded excellent results for the inverse Dirichlet and Neumann problems in the acoustic case. The algorithm is able to return shapes from synthetic data using, respectively, only one or two incident fields. The error is comparable to that occurring using other recently developed methods, but has the great advantage of being able to provide the reconstructions with significantly less data. Results were presented at the APS/URSI International Symposium in Ann Arbor in June 1993 [5.13] and a paper describing

the numerical results is currently in preparation.

A second line of attack on the shape identification problem again involved the use of complete families but instead of simultaneously reconstructing the shape and the solution of a scattering problem for particular boundary data, in this approach we attempt to reconstruct the shape and the Green's function for a given class of boundary conditions, viz. Dirichlet, Neumann or Robin. One advantage of this approach, in contrast to almost all other shape reconstruction methods, is that it readily leads to an algorithm even when scattered field data may only be measured in the backscattering direction. Preliminary results have been reported in [5.11] and [5.16] and a paper describing the method is under preparation. However as yet no numerical experiments have been performed to test the feasibility of the approach.

The third approach we followed concerned the iterative technique developed under AFOSR support that has proven successful in parameter identification problems; specifically reconstructing complex indices of refraction of two dimensional objects from measurements of the fields scattered when the object is illuminated by known sources. Essentially, the method involves casting the problem as an optimization problem in which the cost functional consists of two terms, one is the defect in matching measured data with fields due to a particular index of refraction and the second is the state equation, a set of integral equations in which the index appears and which the fields must satisfy. There are essentially two types of unknown functions, the index of refraction and the total field for each excitation. Each of these functions is constructed iteratively using linear updating, the nonlinear nature of their interrelationship being, nevertheless, retained.

Previous versions of this algorithm led to an empirically determined limit of reconstructibility of $kd|\chi_{max}| \leq 6\pi$ where $k$ is the wave number, $d$ a characteristic diameter of the scatterer, and $|\chi_{max}|$ is the largest contrast that can be reconstructed. No *a priori* information about the scatterer was used. This work is described in [3.8], [3.9], [5.1], [5.2] [5.3] and [5.10].

Under the current grant, however, we exploited the fact that in most problems of interest, the imaginary part of the contrast is non–negative. Incorporating this constraint into the algorithm resulted in a remarkable improvement in the limit of reconstructibility. In fact, using the fact that for extremely good conductors the contrast is essentially large positive imaginary, we successfully reconstructed the boundary of a perfect conductor. In a dramatic demonstration of the efficacy of the method, experimental data provided by Rome Laboratories, Hanscom AFB for a perfectly conducting body was used in our algorithm in a "blind" reconstruction. That is, the actual geometry of the object was not provided, only the experimental scattering data. The algorithm successfully reconstructed the unknown target. Parts of these results have been reported in invited talks in the British Applied Math Colloquium [5.15], the XXIV General Assembly of URSI [5.17] and the Mathematics Forschung Institut, Oberwolfach [5.18]. A paper describing the modifications of the algorithm, *Two Dimensional Location and Shape Reconstruction* [3.10] has appeared in Radio Science and another paper describing the blind reconstruction has been completed [3.13]. The experimental work at Rome Laboratories is being done by Robert McGahan and Marc Coté while the theoretical and numerical work is being done in collaboration with Peter van den Berg of Delft University of Technology in the Netherlands.

## 2.4. Related Work

Work in the three main problem areas described in sections 2.1 – 2.3, was accompanied by some significant related activity which is briefly summarized in this section.

In addition to the applications of optimization methods in antenna problems it was shown that a similar approach could be successfully followed in a class of free surface hydrodynamical problems [3.1], [3.3]. This included development of a constructive method for finding the hull design which optimizes hydrodynamic performance characteristics such as drag and added mass.

The iterative solution of the inverse problems, which has become a major and productive component of the research program, was inspired by previous work on iterative solutions of integral formulations of direct scattering problems. These iterative methods were described in [5.9] and a comprehensive review of these methods in electromagnetics was invited for inclusion in the 1990–1992 Review of Radio Science [3.6]. In addition, uniquely solvable integral equations for electromagnetic scattering from indentations in plane screens were devised [3.7], [5.4]. These equations have application to problems involving small cavities in otherwise smooth surfaces. The subject of small scatterers was also pursued in other ways. Previously obtained results on applications of the Kelvin inversion to low frequency scattering were used to obtain the solution of a canonical low frequency problem, scattering by a concave object [5.7]. The static image theory which attempts to characterize scattered fields by equivalent image sources producing them was extended to the dielectric sphere [3.15], [5.14]. Finally the complete characterization

of the low frequency 'expansion" of the scattered field in two dimensions, when the field is no longer analytic in frequency, was accomplished for arbitrarily shaped scatterers and, in fact, general second order elliptic equations [3.11].

## 3. Publications supported under AFOSR Grant No. 91-0277 (copies included in the Appendix)


1. Recent Developments in Floating Body Problems, T. S. Angell, G. C. Hsaio and R. E. Kleinman, in *Mathematical Approaches in Hydrodynamics*, Touvia Miloh, ed., SIAM Publications, Phila., 141-152, 1991.

2. Antenna Control and Optimization, T. S. Angell, A. Kirsch and R. E. Kleinman, *Proc. IEEE*, 79(1), 1559-1568, 1991 (invited paper).

3. A Constructive Method for Shape Optimization: A Problem in Hydromechanics, T. S. Angell and R. E. Kleinman, *IMA Journ. Appl. Math.*, 47, 265-281, 1991.

4. The Conductive Boundary Condition for Maxwell's Equations, T. S. Angell and A. Kirsch, *SIAM J. Appl. Math*, 52, 1597-1610, 1992.

5. Multicriteria Optimization in Antenna Design, T. S. Angell and A. Kirsch, *Math Methods in the Appl. Sciences*, 15, 647-660, 1992.

6. Iterative Methods for Radio Wave Problems, R. E. Kleinman and P. M. van den Berg, *Review of Radio Science 1990-1992*, W. Ross Stone, ed., Oxford University Press, 1993, 57-74.

7. Electromagnetic Scattering by Indented Screens, J. S. Asvestas and R. E. Kleinman, *IEEE AP* **42**, 22–30, 1994.

8. A Modified Gradient Method for Two-Dimensional Problems in Tomography, R. E. Kleinman and P. M. van den Berg, *J. Comp. and Appl. Math*, 42, 1992, 17-35.

9. An Extended Range Modified Gradient Technique for Profile Inversion, R. E. Kleinman and P. M. van den Berg, *Radio Science*, 28, 1993, 877-884.

10. Two Dimensional Location and Shape Reconstruction, R. E. Kleinman and P. M. van den Berg, *Radio Science* **29**, 1157–1169, 1994.

11. Full Low-Frequency Asymptotic Expansion for Elliptic Equations of Second Order, R. E. Kleinman and B. Vainberg, in *Mathematical and Numerical Aspects of Wave Propagation*, R. Kleinman, et al., eds., SIAM, Philadelphia, PA, 1993, 296-301.

12. Asymptotic Approximation of Optimal Solutions of an Acoustic Radiation Problem, T. S. Angell, R. E. Kleinman, and B. Vainberg, in *Inverse Scattering and Potential Problems in Mathematical Physics*, R. E. Kleinman, R. Kress and E. Martensen, eds. Peter Lang, Frankfurt (in press).

13. Blind Shape Reconstruction from Experimental Data, P. M. van den Berg, M. G. Coté and R. E. Kleinman, submitted to *IEEE-AP*.

14. Modified Green's Functions and Obstacle Reconstruction, R. E. Kleinman and G. F. Roach, in preparation.

15. Low Frequency Image Theory for the Dielectric Sphere, I. V. Lindell, J. C- E. Sten and R. E. Kleinman, *J. Electromagnetic Waves and Applics.*, 8, 295–313, 1994.

16. Full Low–Frequency Asymptotic Expansion for Second–Order Elliptic Equations in Two Dimensions, R. E. Kleinman and B. Vainberg, *Math. Methods in the Appl. Sci.*, (to appear).

## 4. Research Personnel

| | |
|---|---|
| T. E. Angell | – Principal Investigator |
| R. E. Kleinman | – Principal Investigator |
| P. M. van den Berg | – Scientific Investigator |
| Xinming Jiang | – Post Doctoral Investigator |
| Wen Lixin | – Graduate Student |

Note that Dr. Jiang and Ms. Wen received no direct support under the grant but did work on grant related projects.

## 5a. Presentations supported under AFOSR Grant No. 91-0277

1. Two-Dimensional Profile Reconstruction, R. E. Kleinman and P. M. van den Berg, North American Radio Science Meeting, URSI/IEEE-APS, London, Ontario, June 1991.

2. Profile Inversion for Two Dimensional Scatterers, R. E. Kleinman and P. M. van den Berg, PIERS Symposium, Cambridge, MA, July 1991 (invited talk).

3. Iterative Methods for Electromagnetic Profile Inversion, R. E. Kleinman, X. Jiang and P. M. van den Berg, ICIAM, Washington, D.C., July 1991.

4. The Far Field Scattered by Indented Screens, J. S. Asvestas and R. E. Kleinman, National Radio Science Meeting, Boulder, CO, Jan. 1991.

5. A Novel Approach to Antenna Optimization, T. S. Angell, A. Kirsch and R. E. Kleinman, National Radio Science Meeting, Boulder, CO, Jan. 1992.

6. A Rigorous Derivation of Higher Order Boundary Conditions in Electromagnetic Scattering, T. S. Angell, R. E. Kleinman and S. Przeździecki, Wave Phenomena II: Modern Theory and Applications, Edmonton, Alberta, June 1991.

7. Low Frequency Electromagnetic Scattering from Non Convex Bodies, D. Gintides, K. Kiriaki and R. E. Kleinman, IEEE/APS-URSI International Symposium, Chicago, IL, July 1992.

8. Conductive Problems in Scattering with Maxwell's Equation, T. S. Angell, A. Kirsch and R. E. Kleinman, IEEE/APS-URSI International Symposium, Chicago, IL, July 1992.

9. Iterative Methods for Intermediate Frequencies, R. E. Kleinman and P. M. van den Berg, IEEE/APS-URSI International Symposium, Chicago, IL, July 1992 (invited paper).

10. An Extended Range Modified Gradient Technique for Profile Inversion, R. E. Kleinman and P. M. van den Berg, URSI International Symposium on Electromagnetic Theory, Sydney, Australia, August 1992.

11. Obstacle Reconstruction from Back Scattered Data, R. E. Kleinman and G. F. Roach, URSI International Symposium on Electromagnetic Theory, Sydney, Australia, August 1992.

12. Multicriteria Optimization in Arrays, T. S. Angell, R. E. Kleinman and A. Kirsch, *Proceedings of JINA 92 Congress*, Nice, France, 1992.

13. A New Inversion Technique for Shape Reconstruction, T. S. Angell, Xinming Jiang and R. E. Kleinman, URSI Radio Science Meeting, Ann Arbor, MI, June 1993.

14. Low Frequency Image Theory for the Dielectric Sphere, I.V. Lindell and R. E. Kleinman, National Radio Science Meeting, Boulder, CO, Jan. 1992.

15. Modified Gradient Techniques for Profile Inversion, R. E. Kleinman, P. M. van den Berg, British Applied Math Colloquium, Glasgow, Scotland, April, 1993.

16. Modified Green's Functions and Obstacle Reconstruction, G. F. Roach and R. E. Kleinman, British Applied Math Colloquium, Glasgow, Scotland, April 1993.

17. Profile Inversion by Simultaneous Error Reduction, R. E. Kleinman and P. M. van den Berg, XXIVth General Assembly of URSI, Kyoto, Japan, August 1993.

18. Reconstruction of the Location, Shape, and Composition of a Scattering Object, P. M. van den Berg and R. E. Kleinman, Oberwolfach Conference, Methoden und Verfahren der Mathematischen Physik, Dec. 1993.

19. Full Low-Frequency Asymptotic Expansion for Elliptic Equations of Second Order, R. E. Kleinman and B. Vainberg, SIAM-INRIA Conference on Mathematical and Numerical Aspects of Wave Propagation, Newark, DE June 1993.

20. Asymptotic Approximation of Optimal Solutions of an Acoustic Radiation Problem, T. S. Angell, R. E. Kleinman and B. Vainberg, Oberwolfach Conference, Methoden und Verfahren der Mathematischen Physik, December 1993.

21. New Approaches to Numerical Solutions of Integral Equations, R. E. Kleinman, International Conference on Applied and Industrial Mathematics, Linköping, Sweden, June 1994.

## 5b. Interactions with other Laboratories

Rome Laboratory, Hanscom AFB: Collaboration with personnel was initiated and is ongoing in two areas. With R. A. Shore there is a project on applying multicriteria optimization methods to a class of antenna problems previously treated by other methods. With R. V. McGahan and M. G. Coté there is a project devoted to using experimental microwave scattering data as the input in inverse scattering algorithms. This collaboration has resulted in one joint paper submitted for publication and a cooperative effort to organize a workshop on inversion using experimental data.

Laboratory for Electromagnetic Research, Delft University of Technology, The Netherlands: Peter van den Berg of that Laboratory has played a key role in developing robust inversion algorithm and has collaborated on 13 of the papers and presentations listed in Sections 3 and 5a.

Institute for Applied Mathematics, University of Erlangen, Nurenberg, Germany: Andreas Kirsch of the Institute has collaborated on the work on resistive boundary conditions and is currently involved in a joint book project on optimization methods in antenna theory.

Laboratory for Signals and Systems, National Center for Scientific Research (CNRS), France: Collaboration with D. Lesselier and B. Duchene of that Laboratory has begun on extending inversion methods developed under the grant to more complicated problems such as detecting and identifying objects buried in a halfspace from scattering data collected above the half space.

Electronics Laboratory, University of Nice, Sophia Antipolis: Collaboration with Christian Pichot on computing Newton–Kantorovich and modified gradient methods for inverse scattering.

## 6. Discoveries inventions or patent disclosures

There have been no patent applications or inventions under the grant. The research results reported in Sections 3 and 5a are all in their own way discoveries. Perhaps the most striking of these was the success of our modified gradient algorithm in reconstructing the shape and location of an object from experimental scattering data in a "blind" test as described in [3.14].

Appendix: Publications listed in Section 3

In: Mathematical Approaches in Hydrodynamics,
Touvia Miloh, ed., SIAM Publications,
Phila., 1991.

Chapter 10

# Recent Developments in Floating Body Problems

T. S. Angell, G. C. Hsiao, and R. E. Kleinman

## 10.1. Introduction.

In this paper we outline some recent developments in the problem of a body floating on a linearized free surface subject to a time harmonic exciting force. This problem was well known even before Fritz John [11] derived a Green's function satisfying the linearized free surface condition and used that function to prove existence and uniqueness of solutions using integral equation techniques. John followed a standard approach to boundary value problems. First he proved uniqueness, that is there was at most one solution of the boundary value problem. Then he formulated an integral equation whose solutions lead, through an integral representation, to a solution of the boundary value problem, which, since only one was possible, was the unique solution. The existence of a unique solution of the integral equation was established using Fredholm theory. John recognized the existence of "irregular frequencies", discrete real values of the wave number for which this integral equation was not uniquely solvable and he was forced to adopt a more complicated method explicitly involving the eigenfunctions for proving existence and uniqueness for these anomolous values of the wave number. The problem of formulating uniquely solvable integral equations for all frequencies, suitable for numerical solution, has occupied a central position in free surface hydrodynamics for decades. In order to apply the results of potential theory. John made a number of geometric assumptions on the shape of the ship hull. These essentially reduced to requiring that the closed surface formed by a ship hull and its reflection in the free surface formed a smooth (twice differentiable) convex surface. This restriction implies that the ship hull intersects the free surface normally and precludes discontinuities in curvature even on the center plane. Considerable attention has been directed toward relaxing these assumptions.

In this paper we will summarize some recent developments in a number of areas related to the floating body problem. Specifically we will review a number of uniquely solvable integral equations for non–smooth hulls, the present state of attempts to establish the existence of "weak" solutions, and some related optimization problems in hull design.

## 10.2. Notation.

We will concentrate on the three dimensional finite depth case as illustrated in Fig. 1. We fix the origin of a cartesian coordinate system in the water plane that is, the continuation into the ship of the mean free surface so that the $x$–$y$ plane consists of

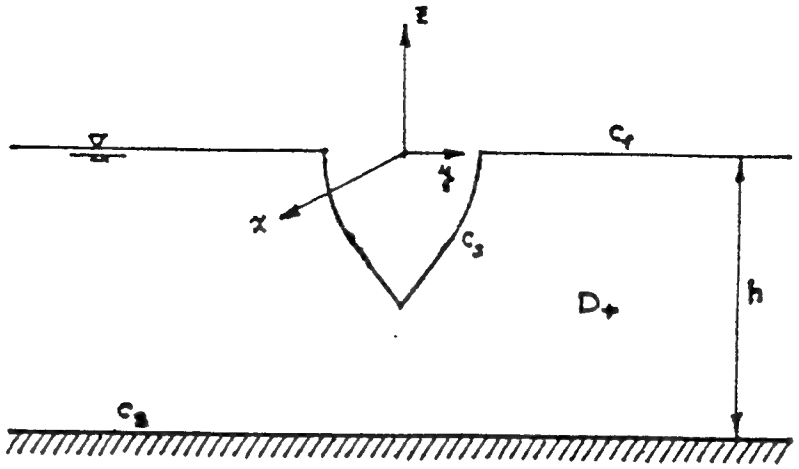



Fig. 1

the mean free surface $c_f$ together with the water plane $c_w$. The fluid domain $D_+$ is bounded by the linearized free surface $c_f$, the wetted portion of the ship hull $c_s$ and the bottom $c_B$ which is assumed flat at a depth $h$, that is, $c_B$ is the plane $z = -h$. Let $D_-$ denote the domain bounded by $c_s$ and $c_w$. We denote position vectors by $\mathrm{p} = (x, y, z)$ and often also employ polar coordinates $x = \rho\cos\theta$, $y = \rho\sin\theta$ in the $x$–$y$ plane.

The time harmonic three dimensional floating body problem with linearized free surface condition is usually formulated in terms of a classical boundary value problem for the complex velocity potential $\phi(\mathrm{p})$ in the fluid domain as follows. Find $\phi$ such that

$$\nabla^2\phi = 0 \text{ in } D_+ , \tag{1a}$$

$$\frac{\partial\phi}{\partial n} = V \text{ on } c_s , \tag{1b}$$

$$\frac{\partial\phi}{\partial n} = 0 \text{ on } c_B , \tag{1c}$$

$$\frac{\partial\phi}{\partial n} + k\phi = -\frac{\partial\phi}{\partial z} + k\phi = 0 \text{ on } c_f , \tag{1d}$$

$$\frac{\partial\phi}{\partial\rho} - i\,k_0\phi = o(\rho^{-1/2}) \text{ as } \rho \to \infty , \tag{1e}$$

where $k$ is the wave number with $\operatorname{Im} k \geq 0$, $k_0$ is the positive real root of

$$k = k_0 \tanh k_0 h , \tag{2}$$

where $\frac{\partial}{\partial n}$ denotes the normal derivative directed into $D_+$. The function $V$ is the specified boundary data on the ship hull and may be chosen so that $\phi$ represents any of the possible radiation components or a diffraction potential. Mathematically it is necessary to specify requirements on $c_s$, the sense in which $\frac{\partial\phi}{\partial n}$ is to be taken on $c_s$, and the class of functions from which the data $V$ is to be chosen in order to have a well posed (that is, uniquely solvable) boundary value problem.

We will denote by $\{v_l\}$ the set of Ursell's multiple potentials for the three dimensional finite depth case (Thorne [25], Martin [19]) which have the property of satisfying

the boundary conditions on $c_f$ and $c_B$. We use the standard multi-index notation

$$l = (n, m, j);\ n, m = 0, 1, 2, \ldots;\ j = 1, 2;\ |l| = n + m + j$$

and define the multipole potentials explicitly to be

$$v_l(\mathrm{p}) := \psi_{nm}(\rho, z)[j \cos m\theta + (1 - j) \sin m\theta]$$

where

$$\psi_{om}(\rho, z) := \int_0^\infty \frac{\cosh \xi(z + h) J_m(\xi\rho)}{\xi \sinh \xi h - k \cosh \xi h}\, d\xi$$

$$\psi_{nm}(\rho, z) := \frac{1}{(2n)!} \int_0^\infty \xi^{m+2n-1}(\xi + k) e^{\xi z} J_m(\xi\rho)\, d\xi$$

$$- \frac{1}{(2n)!} \int_0^\infty \frac{\xi^{m+1} \cosh \xi(z + h) J_m(\xi\rho)}{\xi \sinh \xi h - k \cosh \xi h}\, d\xi\ ,\quad n > 0$$

the contour passing below the singularity $k_0$ in the complex $\xi$-plane. The set $\{v_l\}$ is complete on $L^2(c_s)$ provided $c_s$ satisfies the rather stringent smoothness conditions of John (Martin [20]).

We may now introduce three Green's functions which figure prominently in the integral equation formulation. First is an elementary source which satisfies the boundary condition on $c_B$ but not on $c_f$:

$$\gamma_0(\mathrm{p}, \mathrm{q}) := -\frac{1}{2\pi|\mathrm{p} - \mathrm{q}|} - \frac{1}{2\pi|\mathrm{p} - \bar{\mathrm{q}}|}$$

where

$$\bar{\mathrm{q}} := (x, y, -z - 2h)\ .$$

Next is a Green's function which satisfies the boundary condition on $c_B$ and the free surface condition on $c_f$. This is the Green's function of John for which various representations are known (e.g. Weyhausen and Laitone [28], Noblesse [21]). One such is

$$\gamma_J(\mathrm{p}, \mathrm{q}) := -i \sum_{n=0}^{\infty} \frac{(k_n^2 - k^2) H_0^{(1)}(k_n R)}{h k_n^2 - h k^2 + k} \cosh k_n(z_p + h) \cosh k_n(z_q + h)$$

where

$$R = \sqrt{\rho_p^2 + \rho_q^2 - 2\rho_p \rho_q \cos(\theta_p - \theta_q)}$$

and $k_n$ are the roots of equation (2) with non-negative real and imaginary parts. This Green's function has a source strength double that of $\gamma_0$ on $c_f$ and $c_w$.

Finally we define a modified Green's function (Ursell [26]) to be

$$\gamma_M(\mathrm{p}, \mathrm{q}) := \gamma_J(\mathrm{p}, \mathrm{q}) + \sum_{|l|=0}^{\infty} a_l v_l(\mathrm{p}) v_l(\mathrm{q})$$

where the coefficients $a_l$ are subject to two important restrictions. First the resulting series should be convergent for all $\mathrm{p}, \mathrm{q} \in \overline{D}_+$, e.g. $|a_l| < \frac{c}{n^2 m^2 M_l^2}$ for $n, m > 0$, some $c < \infty$ and $M_l := \sum_{\mathrm{p} \in \overline{D}_+} |v_l|$ and second $\mathrm{Im}\, a_l < 0$ (note that this condition differs from that given by Ursell [26] only because we have chosen a different sign convention for $\gamma_0$ and $\gamma_J$.)

Now we have three Green's functions of ever increasing complexity beginning with the relatively simple $\gamma_0$, the relatively complicated $\gamma_J$ and the even more complicated $\gamma_M$. It should be remarked that all that we are presenting here for the finite depth case may be repeated for infinite depths where the potentials $v_l$ and the Green's functions $\gamma_J$ and $\gamma_M$ simplify considerably.

We can use these Green's functions to define single and double layer potentials as follows: let $\gamma$ denote any one of $\gamma_0, \gamma_J$ or $\gamma_M$ and $c$ denote any one of the surfaces $c_s, c_\infty$, or $c_f$; let $u$ be a function (density) defined on $c$, then define

$$Su := \int_c \gamma(\mathbf{p},\mathbf{q})u(\mathbf{q})\, ds_\mathbf{q} \quad , \quad \mathbf{p} \in D_+ \cup D_- \, ,$$

$$Du := \int_c \frac{\partial \gamma}{\partial n_\mathbf{q}}(\mathbf{p},\mathbf{q})u(\mathbf{q})\, ds_\mathbf{q} \qquad \mathbf{p} \in D_+ \cup D_- \, ,$$

$$\overline{K}^* u := \int_c \frac{\partial \gamma}{\partial n_\mathbf{q}} u(\mathbf{q})\, ds_\mathbf{q} \quad , \quad \mathbf{p} \in c \, ,$$

$$Ku := \int_c \frac{\partial \gamma}{\partial n_\mathbf{p}}(\mathbf{p},\mathbf{q})u(\mathbf{q})\, ds_\mathbf{q} \quad , \quad \mathbf{p} \in c \, ,$$

and

$$Nu := \frac{\partial}{\partial n_\mathbf{p}} \int_c \frac{\partial \gamma}{\partial n_\mathbf{q}}(\mathbf{p},\mathbf{q})u(\mathbf{q})\, ds_\mathbf{q} \quad , \quad \mathbf{p} \in c \, .$$

Note that $\overline{K}^* u$ is the direct value of $Du$ on $c$ and $Su$ may be extended to all points on the closure of $D_+ \cup D_-$. We will append subscripts to indicate the particular choice of $\gamma$ and $c$, e.g.,

$$S_{s,0}u := \int_{c_s} \gamma_0(\mathbf{p},\mathbf{q})u(\mathbf{q})\, ds_\mathbf{q} \, .$$

We remark that the density $u$ and the surface $c$ must be consistently restricted in order for the functions given above to make sense (e.g. Kleinman [12]).

Having established this notation we may now proceed to questions of uniquely solvable integral equations.

### 10.3. Uniquely Solvable Integral Equations.

The uniqueness theorem for the floating body problem is easily proved if Im $k > 0$ for any $c_s$ for which Green's theorem is available in $D_+$., This is shown by John [11] (see also Kleinman [12]). However when Im $k = 0$ additional restrictions on the geometry of $c_s$ are needed, the essential one being that vertical rays from the free surface $c_f$ intersect the ship hull at most once. However it is not necessary to require normal intersections at the free surface (Kuznetsov and Maz'ya [13], Kleinman [12]) nor in fact is it necessary to insist that $c_s$ be smooth. Even with shapes such as those shown in Fig. 2, uniqueness for the floating body problem has been shown provided the angles indicated in Fig. 2 are restricted; $0 < \alpha \leq \pi/2$ and $0 < \beta < 2\pi$. Precise conditions on $c_s$ are given by Kleinman [12] and Wienert [30]. Even the condition that vertical rays from $c_f$ intersect $c_s$ at most once may be relaxed, $\alpha > \pi/2$ in Fig. 2, as was shown by Simon and Ursell [24] in the two–dimensional case. Similar results in three–dimensions are eagerly awaited.

With this uniqueness result for the boundary value problem we know that if an integral equation leads to a representation of a solution of the floating body problem


Fig. 2

it must be *the* solution. In the notation of the previous section a Green's theorem approach leads to a representation of the solution of the floating body problem as

$$\phi := \frac{1}{2}(S_{s,J}V - D_{s,J}w) \quad , \quad \mathrm{p} \in D_+ \, , \tag{3}$$

where $w$ is a solution of the boundary integral equation

$$w + \overline{K}^*_{s,J}w = S_{s,J}V \quad , \quad \mathrm{p} \in c_s \, . \tag{4}$$

Alternatively, using a source distribution approach

$$\phi := S_{s,J}w \quad , \quad \mathrm{p} \in D_+ \, , \tag{5}$$

where $w$ is a solution of

$$w + K_{s,J}w = V \quad , \quad \mathrm{p} \in c_s \, . \tag{6}$$

Note that while we use the same letter to designate the solution of the boundary integral equations in each case, these solutions will be different. In the Green's approach, equation (4), $w$ is the unknown value of $\phi$ on $c_s$ whereas in layer approach it is not.

If Im $k > 0$ then either integral equation is uniquely solvable. However, if Im $k = 0$ there will occur discrete values of $k$ (irregular frequencies) for which there will exist non-trivial solutions of the homogeneous equations

$$w + \overline{K}^*_{s,J}w = 0 \text{ and } w + \overline{K}_{s,J}w = 0 \, .$$

The same values of $k$ will be irregular for both equations but the corresponding solutions (eigenfunctions) will be different.

For Im $k = 0$ we list now a number of boundary integral equations which are uniquely solvable together with the corresponding representation of the velocity potential in the fluid domain.

| | Boundary Integral Equation | Representation of Solution in $D_+$ |
|---|---|---|
| (7) | $w + \overline{K}^*_{s,J} w + \eta N_{s,J} w = S_{s,J} V + \eta(K_{s,J} V - V)$ $\operatorname{Im} \eta \neq 0$ | $\phi = -\frac{1}{2} D_{s,J} w + \frac{1}{2} S_{s,J} V$ |
| (8) | $w + K_{s,J} w + \eta N_{s,J} w = V$ $\operatorname{Im} \eta \neq 0$ | $\phi = S_{s,J} w + \eta D_{s,J} w$ |
| (9) | $w + \overline{K}^*_{s,M} w = S_{s,M} V$ | $\phi = -\frac{1}{2} D_{s,M} w + \frac{1}{2} S_{s,M} V$ |
| (10) | $w + K_{s,M} w = V$ | $\phi = S_{s,M} w$ |
| (11) | $w + \overline{K}^*_{s,J} w + \frac{1}{2} \overline{K}^*_{w,J} w = S_{s,J} V$ | $\phi = -\frac{1}{2} D_{s,J} w + \frac{1}{2} S_{s,J} V$ |
| (12) | $w + \overline{K}^*_{s,0} w + \overline{K}^*_{f,0} w + k S_{f,0} w = S_{s,0} V$ | $\phi = -\frac{1}{2} D_{s,0} w - \frac{1}{2} D_{f,0} w$ $-\frac{k}{2} S_{f,0} w + \frac{1}{2} S_{s,0} V$ |
| (13) | $\int_{c_s} w(q) \frac{\partial}{\partial n_q} v_l(\mathbf{q}) ds_q = \int_{c_s} V(\mathbf{q}) v_l(\mathbf{q}) ds_q$ $\lvert l \rvert = 0, 1, 2, \ldots$ | $\phi = -\frac{1}{2} D_{s,J} w + \frac{1}{2} S_{s,J} V$ |

We assert that all of these integral equations are uniquely solvable however it must be quickly added that rigorous existence proofs have not been carried out in any case. Such proofs are especially difficult for those equations involving the hypersingular operator $N_{s,J}$. Nonetheless existence of a unique solution of a complicated regularized form of (7) has been proven by Wienert [30] for nonsmooth surfaces as depicted in Fig. 2 for all functions $V$ which are integrable on $c_s$. Then, following the procedure described earlier, this produces the solution of the boundary value problem. Once existence of this solution is established, any integral equation derived by a valid use of Green's theorem will be assured of having at least one solution. The only remaining question is uniqueness and this has been established for all of the equations in the table.

Uniqueness for equations (7) and (8) was established by Kleinman [12] (see also Lee and Sclavounos, [15]) and numerical results using these equations have been presented by Lee and Sclavounos [15] and by Lau and Hearn [14]. Note that existence for the layer equation (8) still uncertain.

The modified Green's functions used in equations (9) and (10) stem from the work of Ursell [26] and Martin [19] and Ursell's proof establishes uniqueness. As before, existence for the equation stemming from the use of Green's theorem is assured via the argument based ultimately on Wienert's work, while existence for (10) is uncertain. However, since no hypersingular operators are involved, the theory of boundary integral operators for potential theory in nonsmooth domains (Wendland [29], Burago and Maz'ya [5], Kleinman [12]) should suffice to apply to establish existence and uniqueness for both these equations. The details have not been carried out however. It should be noted that a related modification of the Green's function in which only one multipole potential was added underlay the work of Ogilvie and Shin [22] however while this was shown to eliminate one irregular frequency it will not eliminate all irregular frequencies.

Note that equations (7)–(10) are boundary integral equations on $c_w$ only while equations (11) and (12) are not. Equation (11) is usually attributed to Ohmatsu [23] who in turn credits Wood [31]. Note that the integral equation is over both $c_s$ and $c_w$ while the representation involves the restriction of the solution to $c_s$. Ohmatsu proved uniqueness (see also Kleinman [12], Chang and Pien [6] and, for an alternative

but similar approach, Fernandez, [9]). Equation (11) employs the simplest Green's function at the cost of integrals over the entire free surface $c_f$. Originally derived by Bai and Yeung [4] and Yeung [32], uniqueness was rigorously established by Angell et al. [1]. Numerical work has been done by Bai and Yeung and, for the two dimensional case, by Liu [18].

Finally equation (13) is not strictly a boundary integral equation but an infinite set of moment equations, the so–called null field equations and uniqueness has been shown by Martin [19], [20]. One way of arriving at the null field equations is to employ Green's theorem in $D_-$. Another approach involving this idea supplements the boundary integral equation on $c_s$ with the Green's identity evaluated at specific points in $D_-$. However adding only a finite number of such points does not suffice to establish uniqueness for all $k$. Nevertheless this method originated by Schenck has enjoyed numerical success [14].

### 10.4. Optimal Design.

In this section we turn to certain optimization problems related to the optimal design of the shape of a floating body. Regardless of whether the body is fully or only partially submerged, the quantities of physical interest include not only the wave patterns which can be derived from the velocity potential but also functionals of the potential such as added mass and damping factors which measure the distribution of energy in the fluid (see e.g. Wehausen and Laitone [28], p. 567). These factors are, of course, dependant upon the geometry of the body and it is this dependance which we intend to study in this section. In particular we discuss how these quantities may be optimized over restricted classes of body geometry.

As the quantities to be optimized will depend on the velocity potential which in turn depends on the choice of surface, we will need to consider a *family* of boundary value problems generated by an appropriately chosen collection of possible surfaces. If we denote this collection of surfaces by $\Omega$ then a boundary value problem (1a)–(1e) may be defined for each $c_s \in \Omega$. The fluid domain $D_+$ depends on $c_s$ and is denoted by $D_+(s)$. The corresponding solution may be indexed as $\phi_s$ to explicitly exhibit its dependence on $c_s$. Note that, because we are considering a family of boundary value problems, the data $V$ in (1b) must be defined throughout the domain formed by the union of all surfaces in $\Omega$. This is indeed the case, for example for heaving motion where $V = -\hat{n} \cdot \hat{z}$.

The optimization problems under consideration involve functionals defined on the space of velocity potentials, that is, real numbers associated with each velocity potential which in turn depends on $c_s$ and $V$. When it is convenient to explicitly show this dependence we will write $\phi_s(p, V)$. We confine our discussion to the problem of minimization since maximizing a functional $L$ is equivalent to minimizing $-L$. Thus with $L$ denoting the functional, the optimization problem is that of finding $c_{s_0} \in \Omega$ for a given $V$ such that

$$L(c_{s_0}, V) \leq L(c_s, V) \quad \text{for all } c_s \in \Omega \, .$$

Specific choices of $L$ may be made to embody desirable design criteria. In particular we may choose the functional to be the added mass in which case

$$L(c_s, V) = Re \int_{c_s} \phi_s(c, V) V(\mathbf{p}) ds \, .$$

Note that by appropriate choice of $V$ the functional will represent the added mass associated with any of the six rigid body motions, diffraction or combinations of these motions.

For such optimization problems, there are two basic questions: that of the existence of optimal solutions and, once existence is assured, that of the computation of the optimal solution. We will consider these questions in turn.

In order to show that an optimal solution exists we use a reformulation of the boundary value problem in terms of a uniquely solvable boundary integral equation which will exhibit, explicitly, the dependance of the solution on the surface $c_s$. This is needed in order to establish the continuous dependance of the solution $\phi_s$, not only with respect to the data $V$ which is standard, but also with respect to the surface $c_s$. This is a more difficult mathematical problem. One method of proof paralleling that used by Angell et al. [2] and Angell and Kleinman [3] is to introduce an explicit parameterization of the admissable surfaces and then transform the integral equations on $c_s$ to equations on a reference surface by introducing the Jacobian of the transformation. This essentially moves the dependence on the surface from the solution to the (now more complicated) kernel of the integral equation in which form the continuity is deduced. The technicalities are by no means trivial, even when the admissable surfaces are smooth. For non–smooth surfaces such as those depicted in Fig. 2, the details have yet to appear although the uniquely solvable integral equations of Wienert [30] are thought to provide what is needed for this approach. Once the continuity of the functional $L$ is established on a suitably restricted class of admissable surfaces, it follows that $L$ will assume its absolute minimum on this class .

Turning now to the question of developing a constructive method to approximate the optimal hull configuration, we can extend the method used in [3] to treat the totally submerged body to the more difficult case discussed here. The constructive method proposed in [3] relies, not on the integral equation formulation, but rather on the availability of an appropriate complete family of solutions which we take to be Ursell's multipole potentials $\{v_l\}$ introduced previously. The completness properties, at least for the smooth case, were established by Martin [20] for the two–dimensional infinite depth case. Similar results for the present case remain to be established.

Our construction procedure is a penalization method. Such methods usually involve the introduction of additional terms to the cost functional involving both the partial differential operator and various initial and boundary conditions. The use of complete families allows us to simplify the method by introducing only the penalization term corresponding to the boundary condition on the (unknown) hull.

Specifically we approximate $\phi$ by a linear combination of multipole potentials each of which satisfies all conditions of the boundary value problem (except the boundary condition on $c_s$) for every $c_s \in \Omega$. Then for fixed $N$ we seek constants $\{a_l^{(N)}\}$ and $c_s^{(N)} \in \Omega$ which minimize

$$L_\nu^{(N)}[a_l^{(N)}, c_s^{(N)}] := Re \int_{c_s^{(N)}} \sum_{|l|=0}^{N} a_l^{(N)} v_l(\mathrm{p}) V(\mathrm{p}) ds + \nu \int_{c_s^{(N)}} |\sum_{|l|=0}^{N} a_l^{(N)} \frac{\partial v_l}{\partial n} - V|^2 ds$$

This approach is completly analogous to that introduced for the totally submerged body [3] where we have shown that the sequence of minimizing surfaces $\{c_s^{(N)}\}$ exists and has cluster points and that these cluster points in turn approach a minimizer of $L(c_s, V)$ as $\nu \to \infty$. Details of this approach for the floating body have yet to appear.

### 10.5. Weak Solutions.

In section 1.3 we discussed questions of existence and uniqueness of solutions of the floating body problem, equation (1a–1e), from the viewpoint of boundary integral equations. This indeed is the classical approach to existence for boundary value problems and of course also provides a basis for numerical solution such as the well known

panel methods. However there is another approach which has seen considerable activity in the modern theory of partial differential equations. This involves a Hilbert space or variational approach and leads to so-called weak formulations and corresponding weak solutions which form the basis of finite element methods. While this approach has been used in water wave problems, a complete weak formulation and corresponding existence and uniqueness results for the present problem are still not available. One of the difficulties lies in the fact that the fluid domain is infinite and has an infinite boundary. Some success has been reported by truncating the fluid domain and adapting available results in this simpler context Euvrard et al. [8], Bai and Yeung [4], Yeung [32]. Kuznetsov and Maz'ya [13] considered the entire fluid domain but did not establish existence solely in the weak context, depending instead on the boundary integral results. Another attempt was presented by Lenior and Martin [17] but this result was flawed as discovered by Ursell [27]. In the case when Im $k > 0$, a complete weak approach has been established by Deppel and Hsiao [7]. In this section we give the essentials of this approach and indicate where difficulties remain.

We begin with the derivation of the *sesquilinear form* for the problem. Let us denote by $c_a$ the surface of a cylinder of radius which contains the ship hull ($c_a = \{(\rho, \theta, z) | \rho = a, 0 \le \theta \le 2\pi, 0 < z < -h\}$ and $\rho < a$ for every $\mathbf{p} \in c_s$). Applying Green's theorem to the solution $\phi$ and a test function $\psi$ in the fluid domain within the cylinder $c_a$ leads to

$$\int_{D_+^a} \nabla\phi \cdot \nabla\bar{\psi} d\mathbf{q} - k \int_{c_f^a} \phi\bar{\psi} ds = \int_{c_s} V\bar{\psi} ds + \int_{c_a} \frac{\partial\phi}{\partial n}\bar{\psi} ds$$

where $D_+^a$ and $c_f^a$ denote the portion of $D_+$ and $c_f$ within the cylinder $c_a$. If it is true that

$$\lim_{a\to\infty} \int_{c_a} \frac{\partial\phi}{\partial n}\bar{\psi} ds = 0$$

we arrive at the equation

$$B(\phi, \psi) := \int_{D_+} \nabla\phi \cdot \nabla\bar{\psi} ds_{\mathbf{q}} - k \int_{c_f} \phi\bar{\psi} ds = \int_{c_s} V\bar{\psi} ds \tag{14}$$

which implicitly defines the sesquilinear form $B(\phi, \psi)$. However the radiation condition (1e) assures only that $\phi = O(a^{-1/2})$ on $c_a$ hence the test function $\psi$ must have more rapid decay as $a \to \infty$ in order for the integral over $c_a$ to vanish. Equation (14) provides the basis for the weak formulation only after we specify the function spaces (Hilbert spaces) in which the test functions lie and in which the solution is sought. If the domain were finite then the standard energy space is $H^1$ which essentially includes functions which are square integrable together with their first derivatives over the domain. However for the problem at hand this space $H^1$ is not appropriate for the solution $\phi$ since $\phi$ and its derivatives are not square integrable in $D_+$ unless Im $k > 0$.

For Im $k = 0$ it is necessary to choose a different function space setting in order for equation (14) to be valid. The original boundary value problem dictates properties that the solution space must have in particular regarding the growth as $\rho \to \infty$. Therefore we choose $\phi$ to be in $\tilde{H}^1_{loc}(D_+)$ by which we mean that $\phi$ and its derivatives are square integrable in any finite subdomain of $D_+$ and the restriction of $\phi$ to $c_f$ is square integrable on any finite subdomain of $c_f$.

This is not sufficient to ensure that the radiation condition is fulfilled and therefore this condition must be imposed as an additional restriction. However with the

"solution space" so chosen it is necessary that the test function $\psi$ lie in a different function space. A suitable choice is $H^1_{comp}(\overline{D}_+)$ by which we mean functions which, together with their first derivatives, are square integrable in $D_+$ and in addition each function will vanish outside a bounded subdomain of $\overline{D}_+$ (not necessarily the same subdomain for every function).

Now a precise weak formulation of the floating body problem is the following: given $V \in L_2(c_s)$, find $\phi \in \tilde{H}_{loc}(D_+)$ which also satisfies the radiation condition (1e) such that equation (14) holds for all $\psi \in H^1_{comp}(\overline{D}_+)$. A solution in this case is called a *weak solution* of the floating body problem. The next step is to establish the existence and uniqueness of a weak solution. Although progress in this direction has been made, the complete proof has yet to appear. In order to give some idea of where difficulties remain we will outline an approach when Im $k > 0$ where existence and uniqueness can be established.

When Im $k > 0$ the classical radiation condition implies that solutions decay exponentially as $\rho \to \infty$. This means that $\phi$ and its derivatives are square integrable in $D_+$ and $\phi$ is square integrable on $c_f$. Now we may choose the space of test functions as well as the solution space to be the same namely $H := H_1(D_+) \cap L_2(c_f)$ equipped with a special norm and inner product. We define the inner product on this space to be

$$(\phi, \psi) := \int_{D_+} \nabla\phi \cdot \nabla\overline{\psi} d\mathbf{q} + \int_{c_f} \phi\overline{\psi} ds$$

and the induced norm $||\phi|| = \sqrt{(\phi, \phi)}$. This differs from the standard inner product since there is no term of the form $\int_{D_+} \phi\overline{\psi} d\mathbf{q}$. However it has been shown by Doppel and Hsiao [7] that this term is unnecessary, that is, the norm defined above is equivalent to the standard norm. Moreover the presence of the integral over the free surface enables us to establish the following so-called *coerciveness property* for the sesquilinear form defined in (14). There exists a constant (which may depend on $k$) $\lambda(k) > 0$ such that

$$|B(\phi, \phi)| \geq \lambda(k)||\phi||^2$$

for all $\phi \in H$.

In addition the sesquilinear form is *continuous* in the sense that there is a constant $\alpha$ such that

$$|B(\phi, \psi)| \leq \alpha||\phi||||\psi||$$

for all $\phi, \psi$ in $H$.

Finally the right hand side of equation (14) defines a bounded linear functional on $H$ in the sense that

$$|\int_{c_s} V\overline{\psi} ds| \leq \beta||\psi||$$

for all $\psi \in H$ where $\beta$ is a constant.

The process of establishing these results is not trivial and details may be found in [7]. However once they are established, the existence and uniqueness of a weak solution is a consequence of the important Lax–Milgram lemma, (see e.g. Friedman [10]).

By a weak solution is meant a function $\phi \in H$ such that equation (14) is satisfied for all $\psi \in H$. It is relatively easy to see that this weak solution will also satisfy the requirements given earlier for Im $k = 0$.

In the case when Im $k = 0$, we would hope to establish similar results but the previously mentioned breakdown of square integrability on $D_+$ causes difficulties. It

is possible to prove that the right hand side of equation (14) defines a bounded linear functional on $H^1_{comp}(\overline{D}_+)$ and it is also possible to establish the continuity of the sesquilinear form. However the coerciveness property of the sesquilinear form has not been established and this is a problem of considerable magnitude.

An alternative approach to existence and uniqueness in the case $\mathrm{Im}k = 0$ involves a so–called "limiting absorption" principle see e.g. Leis [16]. This involves taking a limit as $\mathrm{Im}\, k \to 0$ of solutions for $\mathrm{Im}\, k > 0$, whose unique existence is known.

To our best knowledge the only attempt to follow this idea is contained in the work of Lenoir and Martin [17] but their result is subject to question because of the unavailability of a Rellich type lemma in this context.

Acknowledgement: this work was partially supported under AFOSR Grant No. 86–0269.

## 10.6. References.

[1] T.S. Angell, G.C. Hsiao, R.E. Kleinman, *An integral equation for the floating–body problem*, J. Fluid Mech., **166** (1986), pp. 161–171.

[2] T.S. Angell, G.C. Hsiao, R.E. Kleinman, *An optimal design problem for submerged bodies*, Math. Meth. in the Appl. Sci. **8** (1986), pp. 56–76.

[3] T.S. Angell and R.E. Kleinman, *A constructive method for shape optimization: a problem in hydromechanics*, to appear IMA J. Appl. Math.

[4] K.J. Bai and R.W. Yeung, *Numerical solutions to free surface flow problems*, 10th Symposium on Naval Hydrodyn., MIT (1974), pp. 609–647.

[5] Y.D. Burago and V.G. Maz'ya, *Potential theory and function theory for irregular regions*, Seminars in Mathematics. Steklov Inst., Leningrad, 3 (1967) [Consultants Bureau, N.Y. (1969)].

[6] M.S. Chang and P.C. Pien, *Hydrodynamic forces on a body moving beneath a free surface*, Proc. First Int. Conf. on Numer. Ship Hydrodyn., David Taylor NSRDC (1975).

[7] K. Doppel and G.C. Hsiao, *On weak solutions of the floating body problem*, University of Delaware, Department of Mathematical Sciences, Tech. Rep. 87–5 (1987).

[8] D. Euvrard, A. Jami, C. Morice and Y. Dusset, *Calcul numerique des oscillation d'un naivre engendrees par la houle*, I et II, J. Mécanique **16** (1977), pp. 289–394.

[9] A.F. Fernandes, *Automatic adjustments for the removal of irregular frequencies in free–surface two–dimensional problems*, Int. Symp. in Offshore Eng., Rio de Janiero, Brasil, 1989.

[10] A. Friedman, *Partial Differential Equations*, Holt, Rinehart and Winston, New York, 1969.

[11] Fritz John, *On the motion of floating bodies II*, Comm. Pure Appl. Math., 3 (1950), pp. 45–101.

[12] R.E. Kleinman, *On the mathematical theory of the motion of floating bodies - an update*, David W. Taylor Naval Ship Research and Development Center Report DTNSRDC–82/074, Bethesda, MD (1982).

[13] N.G. Kuznetsov and V.G. Maz'ya, *Problems concerning steady–state oscillations of a layer of fluid in the presence of an obstacle*, Sov. Phys. Dokl, 19 (1974), pp. 341–343.

[14] S.M. Lau and G.E. Hearn, *Suppression of irregular frequency effects in fluid–structure interaction problems using a combined boundary integral equation method*, Int. J. for Numer. Meth. in Fluids, 9 (1989), pp. 763–782.

[15] C.-H. Lee and P.D. Sclavounos, *Recovering the irregular frequencies from integral equations in wave-body interactions*, J. Fluid Mech. 207 (1989), pp. 393–418.

[16] R. Leis, *Initial Boundary Value Problems in Mathematical Physics*, John Wiley & Sons, New York, 1986.

[17] M. Lenior and D. Martin, *An application of the principle of limiting absorption to the motions of floating bodies*, J. Math. Anal. Appl. 78 (1981), pp. 370–363.

[18] Y.W. Liu, *A Boundary Integral Method for the Two-dimensional Floating Body Problem.* Ph.D. Dissertation, Department of Mathematical Sciences, University of Delaware, Newark, DE, 1987.

[19] P.A. Martin, *On the null field equations for water-wave radiation problems*, J. Fluid Mech. 113 (1981), pp. 315–332.

[20] P.A. Martin, *On the null-field equations for water-wave scattering problems*, IMA J. Appl. Math., 33 (1984), pp. 55–69.

[21] F. Noblesse, *The Green function in the theory of radiation and diffraction of regular water waves by a body*, J. Eng. Math. 16 (1982), pp. 137–169.

[22] T.F. Ogilvie and Y.S. Shin, *Integral equation solutions for the time dependent free-surface problems*, Naval Arch. and Marine Eng., 16 (1979), pp. 86–96.

[23] S. Ohmatsu, *On the irregular frequencies in the theory of oscillating bodies in a free surface*, Papers of Ship Res. Inst. Tokyo No. 48, (1975).

[24] M.J. Simon and F. Ursell, *Uniqueness in linearized two-dimensional water-wave problems*, J. Fluid. Mech. 148 (1984), pp. 137–154.

[25] R.C. Thorne, *Multipole expansions in the theory of surface waves*, Proc. Camb. Phil. Soc. 49 (1953), pp. 707–716.

[26] F. Ursell, *Irregular frequencies and the motion of floating bodies*, J. Fluid Mech. 105 (1981) pp. 143–156.

[27] F. Ursell, Note on a paper by M. Lenoir and D. Martin, Proc. Symp. on Numer. Ship. Hydrodynamics, Paris, 1981.

[28] J.V. Wehausen and E.V. Laitone, *Surface Waves*, in Handbuch der Physik, Bd. IX: Strömungsmechanik III, S. Flugge ed., Springer–Verlag, Berlin, 1960. pp. 446–978.

[29] W. Wendland, *Die Behandlung von Randwertaufgaben im $R_3$ mit Hilfe von Einfach und Doppelschict Potentialen*, Numer. Math. 11 (1968), pp. 380–404.

[30] L. Wienert, *An existence proof for a boundary-value problem with non-smooth boundary from the theory of water waves*, IMA J. Appl. Math. 40 (1988), pp. 95–112.

[31] P.Wood in State of the Art Report Seakeeping, 16 ATTC, 1972.

[32] R.W. Yeung, *A hybrid integral equation method for time harmonic free surface flows*, Proc. First Int. Conf. on Numer. Ship Hydrodynamics, Gathersburg, MD (1975), pp. 581–607.

# Antenna Control and Optimization

THOMAS S. ANGELL, ANDREAS KIRSCH, AND RALPH E. KLEINMAN

*Invited Paper*

*A class of radiation problems is considered wherein an arbitrary smooth surface on which currents may be induced is treated as an antenna. A variety of measures of antenna performance are defined in terms of functionals of the radiation pattern. These in turn give rise to a class of optimization problems in which the current distribution is sought which maximizes or minimizes one or another of the antenna performance functionals. A general method, based on the use of vector wave functions, of reducing each problem to one in finite dimensions is presented. Some numerical examples are presented to illustrate results attainable by these methods.*

## I. Introduction and Basic Problems

Antennas, devices for transmitting or receiving electromagnetic energy, take on a variety of physical forms. They can be as simple as a single radiating dipole, or far more complicated structures consisting of nets of wires or solid conducting surfaces. In any specific case, questions arise from the desire to control and even optimize the performance of the radiating structure through appropriate "feeding." In this paper, we wish to review some common themes arising in response to these questions and to present a general mathematical formulation which, if not encompassing all such problems, at least may serve as a unifying framework within which we may fruitfully study a significant portion of such applications. In this introductory section we will set notation and formulate some specific radiation problems. The second section is dedicated to a discussion of various measures of antenna performance and the formulation of some typical, and we believe important, optimization problems. The remaining sections are devoted to particular case studies.

The most common class of antenna optimization problems concerns arrays of elementary radiators and the literature abounds in papers dealing with such structures. Here we treat a different class of problems, those involving closed three-dimensional bodies. Rather than consider elementary radiators mounted on conducting bodies we treat the entire body as an antenna and address the question of determining that surface current distribution which optimizes some antenna performance characteristic. We state at the outset that we will make no attempt to discuss practical methods for producing particular current distributions. Such questions are beyond the scope of this paper. Rather, it is our intention to present the analysis of particular mathematical models which can be useful in engineering design, if in no other way, at least in so far as it clarifies theoretical limits.

In this spirit, we will consider a prescribed radiating structure $D_-$, with boundary $S$, as some subset of the usual three-dimensional space, which represents a physical body capable of supporting a flow of electric current. We assume that $D_-$ contains the origin of coordinates. We will also assume that the boundary $S$ is smooth (no edges or corners) and denote the connected exterior region by $D_+$. We denote points by their position vectors $x$ and $y$; if $x \in S$ then the normal to $S$ at the point $x$ will be written $\hat{n}_x$. We will adopt the convention that the unit normal to $S$ at any of its points is directed into the exterior domain $D_-$; the derivative in the direction of $\hat{n}_x$ will be denoted by $\partial/\partial n_x$. We will write $r = |x|$ for the radial variable in spherical coordinates and $\hat{x} = x/r$ as the unit radial vector. In spherical coordinates $\hat{x} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$. Suppose that the region $D_+$ exterior to this body supports an electromagnetic field, denoted as usual by the pair $(E, H)$. Assuming harmonic time dependence $e^{-i\omega t}$, this field is required to satisfy the time–harmonic form of the Maxwell equations

$$\nabla \times E - ikZ_0 H = 0 \tag{1.1a}$$

$$\nabla \times H + ikY_0 E = 0 \tag{1.1b}$$

where $Z_0 = \left(\frac{\mu_0}{\epsilon_0}\right)^{1/2}$, $Y_0 = \frac{1}{Z_0}$, $k = \omega(\epsilon_0\mu_0)^{1/2}$ and $\epsilon_0, \mu_0$ are the free space permittivity and permeability respectively. $Z_0$ and $Y_0$ are the free space impedance and admittance. In the case of solid bodies, the interior may

Manuscript received October 2, 1990; revised February 26, 1991. This work was supported by AFOSR Grant 86-0269 and by NSF Grant DMS-8912593.

T. S. Angell and R. E. Kleinman are with the Department of Mathematical Sciences, University of Delaware Newark, DE 19716.

A. Kirsch is with the Institut für Angewandte Mathematik, Universität Erlangen, Erlangen, Germany.

IEEE Log Number 9103725.

also support such a field (with different $k, Z, Y$) as is the case in the transmission problems arising in electromagnetic scattering theory. We will not discuss these problems here.

We may imagine, in the case of a radiating structure, that the electromagnetic field is produced by a *surface* current, $J$, with total power given by

$$Z_0 \int_S |J|^2 \, ds \tag{1.2}$$

where $|J|$ denotes the magnitude of the complex vector $J$ and $ds$ is the element of surface area (on the unit sphere $ds = \sin\theta \, d\theta \, d\phi$). The surface current $J$ has the same units as $H$ (see (1.5) below) and the factor $Z_0$ ensures that (1.2) has the units of power. Similarly, we may define the total power in the near field in terms of the familiar Poynting vector integrated over a sphere large enough to enclose the antenna. If such a sphere is denoted by $S_a$, then the integral

$$\mathcal{P}_a := \frac{1}{2} \operatorname{Re}\left\{ \int_{S_a} \hat{n} \cdot E \times \bar{H} \, ds \right\} \tag{1.3}$$

represents the power radiated through the sphere $S_a$.

Prescribing the current $J$ on $S$ provides the boundary data required for a well posed radiation problem. Specifically, the radiation problem consists of finding the pair $(E, H)$ such that Maxwell's equations (1.1a)–(1.1b) are satisfied in $D_-$ together with the Silver-Müller radiation condition

$$\begin{aligned} \hat{x} \times \nabla \times E + ikE &= o(1/r) \\ &\text{as} \quad r \to \infty \\ \hat{x} \times \nabla \times H + ikH &= o(1/r) \end{aligned} \tag{1.4}$$

and satisfying the boundary condition

$$\hat{n} \times H = J \quad \text{on} \quad S. \tag{1.5}$$

Calderón [1] has shown that this problem has a unique solution for every $J \in L_t^2(S)$ where $L_t^2(S)$ is the set of all vector functions, defined in $S$, whose normal component vanishes and whose magnitude is square integrable. We consider the problem of finding $J$ so that the unique solution of the radiation problem described by (1.1a), (1.1b), (1.4), and (1.5) behaves optimally with respect to one or another of the criteria described in Section II.

In problems of control of antennas, the property of interest is most often the radiated far field. Recall that the fields $E$ and $H$ are known [4] to have the representation

$$E(x) = \frac{e^{ikr}}{r} F(\hat{x}) + O\left(\frac{1}{r^2}\right), \quad r \to \infty \tag{1.6}$$

$$H(x) = Y_0 \frac{e^{ikr}}{r} \hat{x} \times F(\hat{x}) + O\left(\frac{1}{r^2}\right), \quad r \to \infty. \tag{1.7}$$

The vector function $F$, which has no radial component, is called the radiation pattern. The power radiated into the far field is

$$\mathcal{P} = \frac{1}{2} \operatorname{Re}\left\{ \lim_{r \to \infty} \int_{S_-} \hat{n} \cdot E \times \bar{H} \, ds \right\} \tag{1.8}$$

or in the notation of (1.3), $\mathcal{P} = \lim_{r \to \infty} \mathcal{P}_r$. With (1.6) and (1.7), $\mathcal{P}$ may be written

$$\mathcal{P} = \frac{Y_0}{2} \int_{S_1} |F(\hat{x})|^2 \, ds \tag{1.9}$$

where $S_1$ is the unit sphere. In terms of the $L^2$-norm,

$$\mathcal{P} = \frac{Y_0}{2} \|F\|^2_{L_t^2(S_1)}. \tag{1.10}$$

We use the notation $\|\cdot\|_{L_t^2(S_1)}$ to denote the $L^2$ norm of tangential vectors on $S_1$, the surface of the unit sphere, and $\|\cdot\|_{L_t^2(S)}$ to denote the $L^2$ norm of tangential vectors on the surface $S$. It is important to distinguish between these two norms.

As the problems of antenna control that we intend to discuss are those in which some numerical measure involving the far field is to be optimized by selecting the appropriate surface current $J$ from some preassigned subset of $L_t^2(S)$, it is necessary for us first to understand the mathematical relationship between the current on the radiating body and the far field of the resulting exterior fields. The fact that there exist unique solutions of the exterior boundary value problem guarantees the existence of an operator $\mathcal{K}$ which associates to each admissible current $J$ the corresponding far field $F$. That is, for each admissible $J$ on $S$ there is a unique solution of the boundary value problem, equations (1.1 a,b), (1.4) and (1.5), and this solution, in the far field, has a unique radiation pattern $F$, see equations (1.6) and (1.7), i.e., $\mathcal{K}J = F$. This operator $\mathcal{K}$ is not explicitly known except in special cases but some of its important properties are known. These may be inferred by examining the representation of the fields in terms of dyadic Green's functions [2], [3]:

$$H(x) = \int_S J(y) \cdot \nabla_y \times \bar{\bar{\Gamma}}(x, y) \, ds_y \tag{1.11a}$$

$$E(x) = -\frac{1}{ikY_0} \nabla \times \int_S J(y) \cdot \nabla_y \times \bar{\bar{\Gamma}}(x, y) \, ds_y \tag{1.11b}$$

where $\bar{\bar{\Gamma}}$ is the dyadic

$$\bar{\bar{\Gamma}} = \Gamma_1 \hat{e}_1 + \Gamma_2 \hat{e}_2 + \Gamma_3 \hat{e}_3 \tag{1.12}$$

$\hat{e}_j$ $= 1.2.3$ are rectangular unit vectors, and the vector fields $\Gamma_i$, $i = 1.2.3$, satisfy

$$\nabla_x \times \nabla_x \times \Gamma_i - k^2 \Gamma_i = \hat{e}_i \delta(|x - y|) \tag{1.13a}$$

$$\hat{x} \times \nabla_x \times \Gamma_i + ik\Gamma_i = o(1/r) \quad \text{as} \quad r \to \infty \tag{1.13b}$$

$$\hat{n}_x \times \Gamma_i = 0 \quad \text{on} \quad S. \tag{1.13c}$$

The $\Gamma_i$ have the same asymptotic behavior as the fields $E$ and $H$:

$$\Gamma_i = \frac{e^{ikr}}{r} F_i(\hat{x}, y) + O(1/r^2) \tag{1.14}$$

where the functions $F_i$, $i = 1.2.3.$ are analytic in the sense that their Taylor expansions converge. If we define the dyadic

$$\hat{\bar{\Gamma}}(\hat{x}.y) = \sum_{i=1}^{3} F_i(\hat{x}.y)\hat{e}_i \tag{1.15}$$

then the far field pattern $F$ is given by

$$F(\hat{x}) = -Z_0\hat{x} \times \int_S J(y) \cdot \nabla_y \times \hat{\bar{\Gamma}}(\hat{x}, y)\, ds_y . \tag{1.16}$$

In spite of the fact that the kernel $\bar{\bar{\Gamma}}$ (and hence $\hat{\bar{\Gamma}}$) is not known explicitly except for certain special surfaces $S$, the continuous differentiability of the functions $F_i$ in (1.14) guarantees that the relation (1.16) defines the required operator $\mathcal{K}$. In addition the relation (1.16) ensures that $\mathcal{K}$ is compact, that is, sequences $\{J_n\}$ which are bounded in $L_t^2(S)$ are mapped into sequences $\{\mathcal{K}J_n\}$ which converge to a function in $L_t^2(S_1)$ (more precisely a subsequence will converge). This property is of extreme importance in proving the existence of optimal current distributions. Moreover, by Corollary (4.10) of Colton and Kress [4], $\mathcal{K}$ is a one-to-one mapping. We can also introduce the adjoint operator $\mathcal{K}^*$ (also compact) which associates to each far field $F$ in $L_t^2(S_1)$ a corresponding $J$ in $L_t^2(S)$ through the defining relation

$$(\mathcal{K}J.F)_{L_t^2(S_1)} = (J.\mathcal{K}^*F)_{L_t^2(S)} \quad \text{for all} \quad F \in L_t^2(S_1),\ J \in L_t^2(S). \tag{1.17}$$

where $(\cdot,\cdot)_{L_t^2(S)}$ denotes the inner product in $L_t^2(S)$ and similarly for $S_1$. With this notation we see that we can rewrite (1.9) as

$$\mathcal{P}(J) = \frac{Y_0}{2}(F,F)_{L_t^2(S_1)} = \frac{Y_0}{2}(\mathcal{K}^*\mathcal{K}J, J)_{L_t^2(S)}. \tag{1.18}$$

This form will be particularly useful in our later discussions.

One disadvantage of (1.16) is that, except in particular situations, we have no explicit form for $\hat{\bar{\Gamma}}$ and consequently no simple way to compute the far field generated by a given surface current. Although there are a number of ways attacking this problem numerically, we want a method which will also prove amenable to the optimization problems which are our main concern. One such method involves the use of complete families of solutions whose asymptotic properties are easily calculated. Such families of functions are available in terms of a distribution of dipoles in $D_-$ [1], [5] or alternatively as multipoles at the origin [6]. We choose the latter approach. Following Müller [7] we introduce the functions

$$\begin{aligned} \psi_{n,m} &:= P_n^m(\cos\theta)\cos m\varphi \\ &\quad n = 1,\ldots,\infty, \quad m = 0,\ldots,n \\ \psi_{n,m} &:= P_n^{-m}(\cos\theta)\sin m\varphi \\ &\quad n = 1,\ldots,\infty, \quad m = -1,\ldots,-n \end{aligned} \tag{1.19}$$

where $P_n^m$ are the associated Legendre functions of degree $n$ and order $m$. Denoting the spherical Hankel functions of the first kind by $h_n$, we choose the collection of vector wave functions, defined in $\mathbb{R}^3\backslash\{0\}$,

$$\begin{aligned} \mathcal{G} := \{&\nabla \times (xh_n(kr)\psi_{n,m}), \\ &\nabla \times \nabla \times (xh_n(kr)\psi_{n,m}) : \\ &n = 1,\ldots,\infty,\ m = -n,\ldots,n\} \end{aligned} \tag{1.20}$$

whose far fields form the collection

$$\begin{aligned} \mathcal{F} := \Big\{ &\frac{i^{-n-1}}{k}\nabla\psi_{n,m} \times x, i^{-n}\hat{x} \times (\nabla\psi_{n,m} \times x) : \\ &n = 1,\ldots,\infty,\ m = -n,\ldots,n \Big\}. \end{aligned} \tag{1.21}$$

For simplicity, we will reindex these families and write, simply,

$$\mathcal{G} = \{g_n\}_{n=1}^{\infty}$$

and

$$\mathcal{F} = \{f_n\}_{n=1}^{\infty}.$$

Now define the set of tangential vector fields on the surface $S$ in terms of the restriction of $\mathcal{G}$ to $S$ by

$$\mathcal{G}_t(S) := \{\hat{n} \times g_n\}_{n=1}^{\infty}.$$

Then

$$f_n = \mathcal{K}(\hat{n} \times g_n). \tag{1.22}$$

Since $S$ is taken to be fixed we introduce a shorthand for elements of $\mathcal{G}_t(S)$

$$g_n^t = \hat{n} \times g_n. \tag{1.23}$$

The usefulness of this family is expressed in the following theorem.

*Theorem 1.1.:* Let $S$ be a smooth closed surface containing the origin in its interior. Then the family of functions $\mathcal{G}_t(S)$ is complete and linearly independent in $L_t^2(S)$.

The proof is due to Müller [7] who specifically applied the results of Calderón [1], and Wilcox [5]. Moreover it is straightforward to establish the following.

*Corollary 1.1.:* The family of functions $\mathcal{F}$ is complete and linearly independent in $L_t^2(S_1)$.

Since the set $\mathcal{G}_t(S)$ is complete in $L_t^2(S)$, any surface current can be approximated to any desired degree of accuracy by an appropriate linear combination of elements of $\mathcal{G}_t(S)$. Thus given any $\epsilon > 0$ there exists an integer $N$ and coefficients $c_n^{(N)}$, $n = 1,\ldots,N$ such that

$$\|J - \sum_{n=1}^{N} c_n^{(N)} g_n^t\|_{L_t^2(S)} < \epsilon. \tag{1.24}$$

Here, the choice of coefficients depends on the number of terms used in the approximation for, while the set $\mathcal{G}_t(S)$ is a complete linearly independent set, it is not in general an orthogonal family.

If we write

$$J^{(N)} = \sum_{n=1}^{N} c_n^{(N)} g_n^t \tag{1.25}$$

then the corresponding far field $F^{(N)}$ is simply given by

$$F^{(N)} = \sum_{n=1}^{N} c_n^{(N)} f_n \tag{1.26}$$

the specific form of the $f_n$ being given in (1.21). Certainly, since $\mathcal{K}$ is continuous and bounded, we have that

$$\|\mathcal{K}J^{(N)} - \mathcal{K}J\|_{L_t^2(S_1)} \le c\,\|J^{(N)} - J\|_{L_t^2(S)} \le c\epsilon \tag{1.27}$$

for some suitable constant $c$ so that the far field $F^{(N)}$ likewise approximates $F$ according to (1.27). Thus the use of complete families of solutions will allow us to approximate the far fields produced by given surface currents without the explicit knowledge of the Green's function. We remark that the dependence of the far field on the surface $S$, which is explicit in the use of Green's function, is present in the approximation method in that the coefficients will depend on $S$. It should be noted that these complete families have been used successfully to solve the boundary value problem by minimizing the error in satisfying the boundary condition [8]. However there are many competitive methods, e.g., boundary element or moment methods, which may well be superior for that purpose. Our use of these families is motivated by their convenience in finding approximate solutions of the antenna optimization problems described in the next section rather than solving boundary value problems.

## II. Measures of Antenna Performance

Traditional measures of antenna performance involve a number of scalar quantities including power radiated into the far field as given by (1.9). Other quantities which figure in the literature as useful design parameters include *directivity*, *gain*, and *signal-to-noise ratio (SNR)*. The *directivity* is defined as a normalized ratio of the power radiated in a particular direction to the total power radiated into the far field

$$D(\hat{x}) = 4\pi Y_0 |F(\hat{x})|^2 / \mathcal{P}\,. \tag{2.1}$$

The maximum directivity, i.e., $D := \max_{\hat{x}} D(\hat{x})$ is frequently used as a measure of the ability of the antenna to focus in a given direction. The *gain*, measured with respect to a given direction, is defined as a normalized ratio of the power radiated in a given direction to the total power fed to the antenna. In the present context we make no assumption on the efficiency with which power fed to the antenna is converted into surface current $J$. Rather we define the quantity *radiation efficiency*

$$G(\hat{x}) = 4\pi Y_0^2 |F(\hat{x})|^2 / \|J\|^2_{L_t^2(S)} \tag{2.2}$$

and the corresponding maximal radiation efficiency $G = \max_{\hat{x}} G(\hat{x})$. The quantity $G(\hat{x})$ coincides with the usual concept of gain only if all the power fed to the antenna were converted to surface current. In addition there are various so-called "quality factors," defined differently by various authors, but all intended to measure the "efficiency" of an antenna by comparing the power radiated into the far field and the power supplied to the antenna structure. In particular, one proposal for the quality factor, $Q$, is

$$\begin{aligned} Q := Q(J) &= Z_0^2 \|J\|^2_{L_t^2(S)} / \|F\|^2_{L_t^2(S_1)} \\ &= Z_0^2 \|J\|^2_{L_t^2(S)} / \|\mathcal{K}J\|^2_{L_t^2(S_1)}\,. \end{aligned} \tag{2.3}$$

We will use this form for illustration purposes only and refer to the book of Rhodes [9] for a more complete discussion of quality factors. We pause to remark that this definition of quality factor is connected with the far field operator in a fundamental way. Specifically using the far field operator $\mathcal{K}$

$$\begin{aligned} \|\mathcal{K}\|^2 &= \sup_{J \in L_t^2(S)} \frac{\|\mathcal{K}J\|^2_{L_t^2(S_1)}}{\|J\|^2_{L_t^2(S)}} \\ &= \sup_{J \in L_t^2(S)} \left( \frac{1}{Y_0^2\, Q(J)} \right) > 0\,. \end{aligned} \tag{2.4}$$

Hence,

$$\inf_{J \in L_t^2(S)} Y_0^2\, Q(J) = \frac{1}{\|\mathcal{K}\|^2}\,. \tag{2.5}$$

The notions of directivity and radiation efficiency as given above by (2.1) and (2.2) respectively, represent an idealization of quantities that can actually be measured. It is more realistic to interpret measurements of the intensity in the far field as averages over (perhaps small) patches of the unit sphere. In particular let $\alpha(\hat{x})$ denote the characteristic function of a measurable sector of $S_1$, i.e.:

$$\alpha(\hat{x}) = \begin{cases} 1, & \hat{x} \in \text{sector} \\ 0, & \hat{x} \notin \text{sector.} \end{cases}$$

Then we may generalize the concepts of directivity and radiation efficiency in a particular direction by replacing the expression $|F(\hat{x})|^2$ with an average over a sector containing the particular direction $\hat{x}$:

$$\frac{1}{\|\alpha\|^2_{L^2(S_1)}} \int_{S_1} \alpha(\hat{x})\, |F(\hat{x})|^2\, ds = \frac{\|\alpha F\|^2_{L_t^2(S_1)}}{\|\alpha\|^2_{L^2(S_1)}} \tag{2.6}$$

where $\alpha$ is the characteristic function of the sector. Then

$$D[\alpha] = \frac{4\pi Y_0^2 \|\alpha F\|^2_{L_t^2(S_1)}}{\|\alpha\|^2_{L^2(S_1)}\, \mathcal{P}} \tag{2.7}$$

and

$$G[\alpha] = \frac{4\pi Y_0^2 \|\alpha F\|^2_{L_t^2(S_1)}}{\|\alpha\|^2_{L^2(S_1)} \|J\|^2_{L_t^2(S)}} \tag{2.8}$$

are the generalized directivity and radiation efficiency in the sector characterized by $\alpha$. If the sector shrinks to the point $\hat{x}_0$ ($\alpha(\hat{x}) = 1$, $\hat{x} = \hat{x}_0$; $\alpha(\hat{x}) = 0$ $\hat{x} \neq \hat{x}_0$), then the generalized directivity and radiation efficiency become the pointwise quantities given by (2.1) and (2.2). If the sector is the entire unit sphere ($\alpha_1(\hat{x}) = 1$, $\hat{x} \in S_1$) then

$$D[\alpha_1] = 2 \tag{2.9}$$

and

$$G[\alpha_1] = \frac{1}{Q}. \tag{2.10}$$

Thus the problem of maximizing the reciprocal of $Q$ is a special case of the problem of maximizing the radiation efficiency $G[\alpha_1]$. Corresponding to (2.5) it may be seen that

$$\sup_{J \in L_t^2(S)} G[\alpha] = \frac{4\pi Y_0^2 \|\alpha \mathcal{K}\|^2}{\|\alpha\|^2_{L_t^2(S_1)}}. \tag{2.11}$$

Another performance criterion involves the concept of noise. Noise may be characterized by a function $\omega$ defined on the unit sphere which distorts the radiation pattern. A measure of how much the radiated field is distorted by noise is given by

$$\int_{S_1} \omega(\hat{x})^2 \, |F(\hat{x})|^2 \, ds$$

and the SNR is defined to be

$$\mathrm{SNR}\,(\hat{x}) := \frac{|F(\hat{x})|^2}{\int_{S_1} \omega(\hat{y})^2 \, |F(\hat{y})|^2 \, ds} \tag{2.12}$$

where the variables of integration in the denominator have been changed to avoid confusion. Again, we may define a generalized SNR as

$$\mathrm{SNR}\,[\alpha] := \frac{\|\alpha F\|^2_{L_t^2(S_1)}}{\|\alpha\|^2_{L^2(S_1)} \int_{S_1} \omega(\hat{x})^2 \, |F(\hat{x})|^2 \, ds}. \tag{2.13}$$

Since we can relate the far field to the surface current which generates it, then problems of optimizing a specific performance criterion can be viewed as particular cases of the problem of maximizing a functional $\mathcal{J}(J)$ where the surface current $J$ varies in some set of functions on the surface $S$ which represents physically realizable currents. This requirement of physical realizability typically introduces constraints in the optimization problem. If $U$ represents the set of realizable currents the constrained optimization problem is that of finding

$$\sup_{J \in U} \mathcal{J}(J). \tag{2.14}$$

The question of a concrete description of this set $U$ is an engineering problem the details of which are beyond the scope of the present discussion. One may refer to [10], [11] for typical examples. For *mathematical* reasons it is necessary to require that $U$ have certain properties, namely that it be closed, bounded, and convex, i.e., that it contains all line segments joining two points of $U$. In our examples we take only the most obvious constraint sets. In particular, we constrain the total power on the antenna by taking the bound

$$\|J\|_{L_t^2(S)} = \left( \int_S |J(x)|^2 \, ds \right)^{1/2} \leq M \tag{2.15}$$

for a suitable constant $M$. Such a constraint is intuitively reasonable as it is certainly true that the power supplied to an antenna is limited.

In addition, such quantities as the directivity or the quality factor may well figure in the definition of admissible inputs. Thus for example, a bound on the quality factor limits the amounts of energy stored in the near field. The relation (2.5) shows that $Y_0^2 Q$ is bounded by $1/\|\mathcal{K}\|^2$ where $\mathcal{K}$ is the far field operator. Insofar as the far-field operator is determined by the physical structure of the antenna, e.g., its shape, or the materials from which it is made, we can say that antenna structures with large values of $\|\mathcal{K}\|$ have low quality factors and hence store less energy in the near field.

With these various measures of antenna performance in mind, a number of specific optimization problems may be formulated. Roughly speaking, the optimization problems in the literature fall into two categories. The first, which we may call the synthesis approach, specifies a desired far field pattern (which may or may not be realizable) and asks for an admissible current that will produce a far field most closely approximating the desired pattern. The second, more indirect approach, chooses some performance criterion associated intrinsically with the antenna and asks for that current which optimizes that criterion, as for example, when one asks to maximize the radiation efficiency. For the purposes of the present discussion we classify some common unconstrained problems in the Table 1.

All of the optimization problems listed in the third column are of the form: find $J$, the surface current in $U$, the class of admissible currents, which optimizes a cost functional. The cost functional (e.g., $\int_{S_1} |\mathcal{K}J - \tilde{F}|^2 \, ds$, $\int_{S_1} \alpha(\hat{x}) \, |\mathcal{K}J(\hat{x})|^2 \, ds$, etc.) associates a nonnegative real number with each $J$ in $U$ and the optimization problem is to find that $J$ which produces the largest or smallest such number.

It is often the case that one performance criterion is used as a constraint in the optimization problem for one of the other criteria. Thus for example Kirsch and Wilde [12], [13] have considered the problem of maximizing the SNR subject to an equality constraint on the quality factor. Moreover, certain of these criteria, as with Kirsch and Wilde, are chosen so that the optimization implicitly controls the occurrence of undesirable effects. Clearly, maximizing the SNR inherently constrains the noise. Similarly maximizing radiation efficiency involves a balance between the far field in the desired direction and the amount of power at the radiating surface.

Another approach is to use constraints which explicitly restrict inputs or outputs. Consider, for example the generalized SNR with $\omega$ chosen to be $1 - \alpha$. Then we have a criterion which tends to increase the power in the desired sector while decreasing the power in the complement. This can be viewed as one approach to maximizing power in the beam while keeping power low in the side lobes. Alternately, one could proceed as suggested by Angell and Kleinman [14], by maximizing the functional

$$\int_{S_1} \alpha \, |F(x)|^2 \, dx \tag{2.17}$$

Table 1

| Performance Criterion | Definition | Optimization Problem |
|---|---|---|
| Pattern matching to desired $\tilde{F}$ (continuous) | $\int_{S_1} \lvert F(\hat{x}) - \tilde{F}(\hat{x})\rvert^2 ds$ | $\min_{J\in U} \int_{S_1} \lvert \mathcal{K}J - \tilde{F}\rvert^2 ds$ |
| Pattern matching to desired $\tilde{F}$ (discrete) | $\sum_{i=1}^{N} \lvert F(\hat{x}_i) - \tilde{F}(\hat{x}_i)\rvert^2$ | $\min_{J\in U} \sum_{i=1}^{N} \lvert \mathcal{K}J(x_i) - \tilde{F}(\hat{x}_i)\rvert^2$ |
| Power in a sector with characteristic $\alpha$ | $\frac{Y_0}{2} \int_{S_1} \alpha(\hat{x}) \lvert F(\hat{x})\rvert^2 ds$ | $\max_{J\in U} \int_{S_1} \alpha(\hat{x}) \lvert \mathcal{K}J(\hat{x})\rvert^2 ds$ |
| Signal to noise ratio | $\frac{\lvert F(\hat{x})\rvert^2}{\int_{S_1} \omega(\hat{y})^2 \lvert F(\hat{y})\rvert^2 ds}$ | $\max_{J\in U} \frac{\lvert \mathcal{K}J(\hat{x})\rvert^2}{\int_{S_1} \omega(\hat{y})^2 \lvert \mathcal{K}J(\hat{y})\rvert^2 ds}$ |
| Generalized signal to noise ratio | $\frac{\int_{S_1} \alpha(\hat{x}) \lvert F(\hat{x})\rvert^2 ds}{\int_{S_1} \alpha\, ds \int_{S_1} \omega(\hat{x})^2 \lvert F(\hat{x})\rvert^2 ds}$ | $\max_{J\in U} \frac{\int_{S_1} \alpha(\hat{x}) \lvert \mathcal{K}J(\hat{x})\rvert^2 ds}{\int_{S_1} \alpha\, ds \int_{S_1} \omega(\hat{x})^2 \lvert \mathcal{K}J(\hat{x})\rvert^2 ds}$ |
| Quality factor | $\frac{Z_0^2 \int_S \lvert J\rvert^2 ds}{\int_{S_1} \lvert F\rvert^2 ds}$ | $\min_{J\in U} \frac{\int_S \lvert J\rvert^2 ds}{\int_{S_1} \lvert \mathcal{K}J\rvert^2 ds}$ |
| Radiation efficiency | $\frac{4\pi Y_0^2 \lvert F(\hat{x})\rvert^2}{\int_S \lvert J\rvert^2 ds}$ | $\max_{J\in U} \frac{4\pi \lvert \mathcal{K}J(\hat{x})\rvert^2}{\int_S \lvert J\rvert^2 ds}$ |
| Generalized radiation efficiency | $\frac{4\pi Y_0^2 \int_{S_1} \alpha \lvert F\rvert^2 ds}{\int_{S_1} \alpha\, ds \int_S \lvert J\rvert^2 ds}$ | $\max_{J\in U} \frac{4\pi \int_{S_1} \alpha \lvert \mathcal{K}J\rvert^2 ds}{\int_{S_1} \alpha\, ds \int_S \lvert J\rvert^2 ds}$ |
| Directivity | $\frac{4\pi Y_0 \lvert F(\hat{x})\rvert^2}{\int_{S_1} \lvert F\rvert^2 ds}$ | $\max_{J\in U} \frac{4\pi \lvert \mathcal{K}J(\hat{x})\rvert^2}{\int_{S_1} \lvert \mathcal{K}J\rvert^2 ds}$ |
| Generalized directivity | $\frac{4\pi Y_0 \int_{S_1} \alpha \lvert F\rvert^2 ds}{\int_{S_1} \alpha\, ds \int_{S_1} \lvert F\rvert^2 ds}$ | $\max_{J\in U} \frac{4\pi \int_{S_1} \alpha \lvert \mathcal{K}J\rvert^2 ds}{\int_{S_1} \alpha\, ds \int_{S_1} \lvert \mathcal{K}J\rvert^2 ds}$ |

with the side condition

$$\int_{S_1} \left|(1-\alpha(\hat{x}))\,|F(\hat{x})|^2\,\right| \le M \tag{2.18}$$

for some suitable choice of constant $M$. It would be interesting to compare the results of these two approaches.

## III. Maximizing Power Radiated in a Sector

We illustrate this general approach to antenna optimization problems with one of the specific problems discussed in the previous section, namely determining the current distribution on a surface $S$ which optimizes the power radiated in a sector of $S_1$ with characteristic function $\alpha$. This is done by reducing the problem to a generalized eigenvalue problem which is solved approximately by projection onto a finite dimensional subspace. Here we will summarize the detailed results of [14], [15], and [16].

As indicated in the table of Section II the optimization problem is to find

$$\max_{J\in U} \int_{S_1} \alpha(\hat{x})\,|\mathcal{K}J(\hat{x})|^2\,ds$$

where, using the constraint given in (2.15) with $M = 1$,

$$U = \{J \in L_t^2(S) \mid \|J\|_{L_t^2(S)} \le 1\}. \tag{3.1}$$

Since $\alpha(\hat{x})$ is real ($\alpha = 1$ or 0) we may generalize (1.19) and write

$$\begin{aligned}\mathcal{P}(J;\alpha) &= \int_{S_1} \alpha(\hat{x})\,|\mathcal{K}J|^2\,ds \\ &= (\alpha\mathcal{K}J, \mathcal{K}J)_{L_t^2(S_1)} = (\mathcal{K}^*\alpha\mathcal{K}J, J)_{L_t^2(S)}.\end{aligned} \tag{3.2}$$

Since a current $J$ which optimizes $\mathcal{P}(J;\alpha)$ also optimizes any constant times $\mathcal{P}(J;\alpha)$, for notational convenience we omit the factor $\frac{Y_0}{2}$ in the definition of $\mathcal{P}(J;\alpha)$. We recall that the operator $\mathcal{K}$ which maps $L_t^2(S)$ into $L_t^2(S_1)$ and its adjoint $\mathcal{K}^*$ which maps $L_t^2(S_1) \to L_t^2(S)$ are compact but not in general known explicitly. Even so this characterization of the cost functional $\mathcal{P}(J,\alpha)$ proves very useful in determining approximate optimizers.

The first question to be answered, however, is whether the optimization problem has a solution, that is, does there exist $J_0 \in U$ such that

$$\mathcal{P}(J,\alpha) \le \mathcal{P}(J_0,\alpha) \quad \text{for all} \quad J \in U. \tag{3.3}$$

The answer is in the affirmative and moreover there is an optimizer $J_0$ for which $\|J_0\|_{L_t^2(S)} = 1$ even though we search in $U$ which contains surface currents for which $\|J\|_{L_t^2(S)} \le 1$. Details may be found in [14].

Having established existence of an optimal current it remains to actually find it. To this end the characterization of the cost functional in terms of the operators $\mathcal{K}$ and $\mathcal{K}^*$ in (3.2) proves useful. First observe that the operator

$$R := \mathcal{K}^*\alpha\mathcal{K} : L_t^2(S) \to L_t^2(S) \tag{3.5}$$

is self-adjoint, compact, and non-negative since

$$\begin{aligned}(\mathcal{K}^*\alpha\mathcal{K}J, J)_{L_t^2(S)} &= (\alpha\mathcal{K}J, \mathcal{K}J)_{L_t^2(S_1)} \\ &= \|\alpha\mathcal{K}J\|^2_{L_t^2(S_1)} \ge 0.\end{aligned} \tag{3.6}$$

This means that the spectrum of $R$ is discrete, real and nonnegative and the multiplicity of all nonzero eigenvalues is finite. It then follows that if $(\lambda_0, J_0)$ is an eigenvalue–eigenfunction pair for $R$ such that $\lambda_0$ is the largest

eigenvalue

$$\sup_{J\in\mathcal{U}_1}(RJ,J)_{L_t^2(S)}=(RJ_0,J_0)_{L_t^2(s)}=\lambda_0\,. \tag{3.7}$$

This reduces the optimization problem to an eigenvalue problem.

The next step is to provide a method for approximating the optimizer $J_0$, the associated far field $F_0$ and the optimal power $\mathcal{P}(J_0,\alpha)=\lambda_0$. Here we make use of the complete family $\mathcal{G}_t(S)$ introduced in Section I. Our strategy is to employ a Galerkin procedure which involves projections onto finite dimensional subspaces. Let us define the finite dimensional spaces

$$\mathcal{G}_t^{(N)}=\text{span}\,\{g_n^t\}_{n=1}^N \tag{3.8}$$

and the associated far fields

$$F^{(N)}=\text{span}\,\{f_n=\mathcal{K}g_n^t\}_{n=1}^N \tag{3.9}$$

where $f_n$ and $g_n^t$ are given explicitly in (1.19)–(1.23). In the finite dimensional space the optimal current is of the form

$$J=\sum_{n=1}^N C_n g_n^t \tag{3.10}$$

subject to the constraint

$$\|J\|_{L_t^2(S)}=\left\|\sum_{n=1}^N C_n g_n^t\right\|_{L_t^2(S)}=1\,.$$

The eigenvalue equation

$$RJ=\lambda J \tag{3.11}$$

becomes

$$\sum_{n=1}^N C_n\,R g_n^t=\lambda^{(N)}\sum_{n=1}^N C_n\,g_n^t \tag{3.12}$$

and we seek the largest eigenvalue which we denote by $\lambda_0^{(N)}$. Forming the inner product with $g_m^t$ leads to the generalized algebraic eigenvalue problem

$$\sum_{n=1}^N C_n(Rg_n^t,g_m^t)_{L_t^2(S)}=\lambda^{(N)}\sum_{n=1}^N C_n(g_n^t,g_m^t)_{L_t^2(S)}\,. \tag{3.13}$$

But

$$\begin{aligned}(Rg_n^t,g_m^t)_{L_t^2(S)}&=(\mathcal{K}^*\alpha\mathcal{K}g_n^t,g_m^t)_{L_t^2(S)}\\&=(\alpha\mathcal{K}g_n^t,\mathcal{K}g_m^t)_{L_t^2(S_1)}\\&=(\alpha f_n,f_m)_{L_t^2(S_1)}\,.\end{aligned} \tag{3.14}$$

Here we see the value of the particular family $\mathcal{G}_t(S)$ because even though the operator $R$ may not be explicitly known, the functions $g_n^t$ and $f_n$ are explicitly known (see (1.20) and (1.21)) hence the quantities $(g_n^t,g_n^t)_{L_t^2(S)}$ and $(\alpha f_n,f_m)_{L_t^2(S_1)}$ may be calculated explicitly. The problem of antenna optimization is thereby reduced to determining the largest eigenvalue $\lambda_0^{(N)}$ and an associated $N$-component eigenvector $(C_{0,n})$ (there may be more than one) of the generalized algebraic eigenvalue equation

$$\sum_{n=1}^N C_n(\alpha f_n,f_m)_{L_t^2(S_1)}=\lambda^{(N)}\sum_{n=1}^N C_n(g_n^t,g_m^t)_{L_t^2(S)}\,. \tag{3.15}$$

The fact that the solution of this finite dimensional problem can be used to approximate the solution the original optimization problem is considered next. Define the functions

$$J_0^{(N)}:=\sum_{n=1}^N C_{0,n}\,g_n^t \tag{3.16}$$

and

$$F_0^{(N)}:=\sum_{n=1}^N C_{0,n}\,f_n\,. \tag{3.17}$$

Since we have not imposed the requirement that $\|J_0^{(N)}\|_{L_t^2(S)}=1$, let us introduce normalized coefficients

$$\hat{C}_{0.n}=\frac{C_{0.n}}{\|J_0^{(N)}\|_{L_t^2(S)}} \tag{3.18}$$

and define

$$\hat{J}_0^{(N)}:=\sum_{n=1}^N \hat{C}_{0.n}\,g_n^t \tag{3.19}$$

and

$$\hat{F}_0^{(N)}:=\sum_{n=1}^N \hat{C}_{0.n}\,f_n\,. \tag{3.20}$$

It then follows, as proven explicitly in [15], that

$$\lim_{N\to\infty}\lambda_0^{(N)}=\lambda_0\,. \tag{3.21}$$

Moreover there exists a subsequence of $\{\hat{J}_0^{(N)}\}_{N=1}^\infty$, call it $\{\hat{J}_0^{(M)}\}_{M=1}^\infty$, such that

$$\lim_{M\to\infty}\hat{J}_0^{(M)}=J_0 \tag{3.22}$$

where $J_0$ is an optimal current distribution, i.e.,

$$\lambda_0=\mathcal{P}(J_0;\alpha)\,, \tag{3.23}$$

and an optimal radiation pattern is given by

$$F_0=\lim_{M\to\infty}\hat{F}_0^{(M)}\,. \tag{3.24}$$

This procedure has been carried out numerically for spherical and ellipsoidal antennas [17]. As an illustration we present in Fig. 1 the optimal radiation patterns for three different surfaces, a sphere of radius $a$, ellipsoid with semi axes $.9a, a, 1.1a$ and ellipsoid with semi axes $.5a, a, 1.5a$. Our examples are chosen to demonstrate the effectiveness of this optimization method in finding current distributions which give rise to radiation patterns with two separated

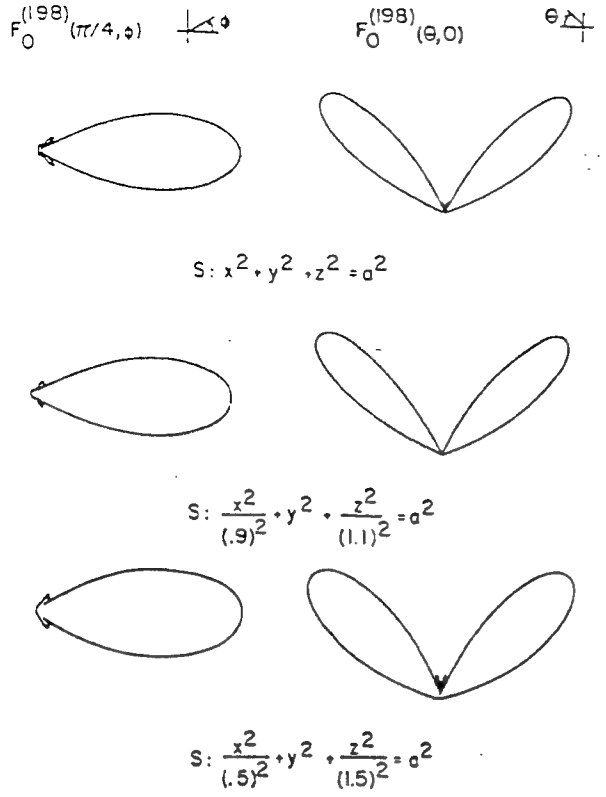



**Fig. 1.** Maximizing Power Radiated in a Sector-Magnitude of optimal radiation patterns for ellipsoidal antennas with two main lobes.

main lobes. Thus the sector in which power is to be maximized was characterized by the function

$$\alpha = \begin{cases} 1, & |\theta - \frac{\pi}{4}| \le \frac{\pi}{128},\ |\phi| \le \frac{\pi}{128} \\ 1, & |\theta - \frac{3\pi}{4}| \le \frac{\pi}{128},\ |\phi| \le \frac{\pi}{128} \\ 0. & \text{all other } (\theta,\phi) \end{cases} \tag{3.25}$$

The magnitude of the optimal radiation pattern is plotted in Fig. 1 for two planar sections, one longitudinal bisecting both main lobes and one latitudinal containing one main lobe. The same sections are used for all three surfaces and the shape of the optimal radiation patterns is seen to be remarkably similar, although it should be remarked that quantitative comparisons are not to be made since different scales were used in different patterns. All calculations were carried out for $ka = 10$ and $N = 198$ (corresponding to a maximum order of 9 for the spherical Hankel functions in (1.20)). We note that while these patterns display main beams which are considerably wider than the sectors characterized by $\alpha$, the qualitative structure of the patterns conforms well with the desired performance criterion embodied in the optimization problem. This example illustrates the feasibility of actually solving one of the optimization problems of Section II and thus provides the antenna designer with a usable mathematical tool for finding optimal current distributions on conformal antennas of known configuration.

For a different approach to this problem of maximizing radiated power, we refer the reader to the Diploma Thesis of H. -G. Burdinsky [18] who uses a gradient method to approximate the optimal solutions.

## IV. Optimization of SNR

As a second example, we consider the problem of optimizing the SNR subject to a constraint on the $Q$-factor (see (2.3)). This example is an extension to the three-dimensional case for Maxwell's equation of the two dimensional problem considered in [12], [13]. Thus we consider here the problem of maximizing the functional

$$\mathrm{SNR}(\hat{x}) = \frac{|F(\hat{x})|^2}{\int_{S_1} \omega(\hat{y})^2 |F(\hat{y})|^2\, ds} = \frac{|\mathcal{K}J(\hat{x})|^2}{\int_{S_1} \omega(\hat{y})^2 |\mathcal{K}J(\hat{y})|^2\, ds} \tag{4.1}$$

subject to a constraint

$$Y_0^2\, Q(J) = \|J\|^2_{L_t^2(S)} / \|\mathcal{K}J\|^2_{L_t^2(S_1)} \le C \tag{4.2}$$

where, in (4.1), $\omega$ represents the noise distribution which is assumed to be nonzero on at least some portion of $S_1$ while, in (4.2) $C$ is a given constant.

For each fixed value of $\hat{x}$, (4.1) defines a functional of $J$ which we denote by $\mathrm{SNR}(J)$. Hopefully this abuse of notation will cause no confusion. The denominator of this functional vanishes only if the function $\omega(\hat{x})F(\hat{x}) = 0$ almost everywhere however our assumptions prevent this from occurring.

The basic existence result of [13] states that if there is any nontrivial $J \in L_t^2(S)$ satisfying the constraint (4.2) then there exists an optimal solution, that is, if $v_0 = \sup\{\mathrm{SNR}(J) \mid J \ne 0,\ Q(J) \le C\}$, then $v_0$ is finite and there exists some admissible $J_0$ such that $\mathrm{SNR}(J_0) = v_0$. The proof relies heavily on the fact that not only is $\mathcal{K}$ a compact mapping from square integrable functions on $S$ to *square integrable* functions on $S_1$ but $\mathcal{K}$ is also bounded as a map onto *continuous* functions on $S$. The details of the proof for the electromagnetic case discussed here may be inferred from the proof for the two-dimensional case appearing in [13]. Note that it is always possible to ensure the existence of $J$ satisfying the constraint (4.2) by taking $C$ sufficiently large.

Further analysis shows that, at each optimal solution, $J_0$, of this problem the constraint (4.2) is *active* by which is meant that

$$Y_0^2\, Q(J_0) = C \tag{4.3}$$

(see [13] Theorem 2.2 for details). Clearly under these circumstances any solution of the optimization problem is likewise a solution of the related problem of maximizing the functional (4.1) subject to the equality constraint (4.3).

For numerical calculations we replace the Hilbert space $L_t^2(S)$ by the finite dimensional space $\mathcal{G}_t^{(N)}$ (see (3.8)), which then replaces the original constrained optimization problem with one in finite dimensions:

$$\max \mathrm{SNR}(J) \tag{4.4}$$

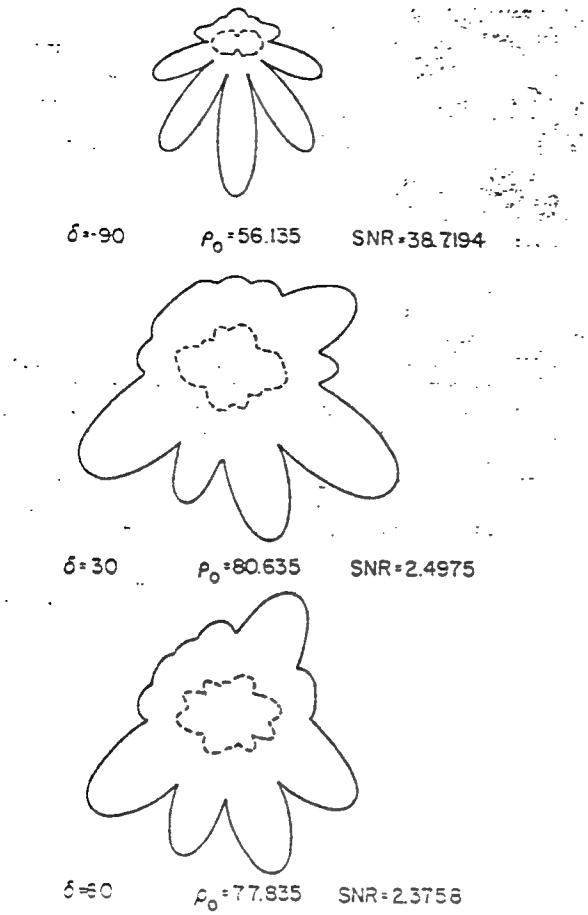



Fig. 2. Maximizing SNR-magnitude of optimal far field (—) and the density of the associated single layer (- - -).

for

$$J \in \mathcal{G}_t^{(N)} \backslash \{0\} \tag{4.5}$$

subject to

$$Y_0^2\, Q(J) \le C. \tag{4.6}$$

Choosing $C \ge \dfrac{\|g_1^t\|^2_{L_t^2(S)}}{\|f_1\|^2_{L_t^2(S_1)}}$ ensures the solvability of the finite dimensional optimization problems for all $N$. In order for the solutions of these finite dimensional problems to be useful, we must, of course, have a convergence result. Such a result is given by Theorem 3.2 of [13]. In the present context we can assert that the set $\{J_0^{(N)}\}$ of normalized optimal solutions of the finite dimensional problems has at least one accumulation point in $L_t^2(S)$ and every such accumulation point is optimal for the original problem.

Actual computations for the finite dimensional problems are carried out in [12] using a generalized cost functional in which the quality factor constraint is included with a Lagrange multiplier. The multiplier rule, applied in the space $L_t^2(S)$, leads to a system of nonlinear equations for the multiplier $\rho_0$ and the optimal current $J_0$,

$$\frac{1}{a}\int_S J_0(y_1)\,\Phi(\hat{x}, y_1)\, ds_{y_1} \cdot \bar{\Phi}^T(\hat{x}, y)$$

$$- \frac{1}{a^2}\,|\mathcal{K} J_0(\hat{x})|^2\, \mathcal{K}^* \omega^2 \mathcal{K} J_0$$
$$+ \rho_0 (I - C\mathcal{K}^*\mathcal{K})\, J_0 = 0 \tag{4.7}$$

$$(J_0, (I - C\mathcal{K}^*\mathcal{K})\, J_0) = 0 \tag{4.8}$$

where the tensor $\Phi(x, y) = Z_0 \nabla_y \times \hat{\Gamma}(x, y) \times x$, which are then projected into the finite dimensional space $\mathcal{G}_t^{(N)}$.

Actually, computations were carried out in [12] for the two dimensional case for H-polarization ($J = \hat{t} u(x, y)$ on $S$ where $\hat{t}$ is a unit tangent vector) for $u$ of the special form of a single layer distribution with density $h$. The surface $S$ was taken to be a circle of radius a and the noise distribution $\omega$ was taken to be the characteristic function of an arc of 200°, that is for

$$\hat{y} = (\cos\theta,\ \sin\theta),\ -\pi \le \theta < \pi$$
$$\omega(\hat{y}) = \begin{cases} 1, & |\theta - \frac{\pi}{2}| \le 100° \\ 0, & |\theta - \frac{\pi}{2}| > 100°. \end{cases}$$

Figure 2 shows the magnitudes of the optimal far field pattern and the density of the single layer which produces this far field for three choices of the direction $\hat{x} = (\cos\delta,\ \sin\delta)$ in which SNR $(\hat{x})$ is maximized. Also included are the optimal values of SNR and the multiplier $\rho_0$. In these examples $ka = 6$, $C = 10$, $N = 15$.

## References


[1] A. Calderón, "Multipole expansion of radiation fields," *J. Rat. Mech. Anal.* vol. 3, pp. 523–537, 1954.

[2] D. S. Jones, *Theory of Electromagnetism*, Oxford, England: Pergamon Press Ltd., 1964.

[3] C. Tai, *Dyadic Green's Functions in Electromagnetic Theory*, Scranton, PA: Intext Educational Publishers, 1971.

[4] D. L. Colton and R. Kress, *Integral Equation Methods in Scattering Theory*, New York, NY: Wiley Interscience, 1983.

[5] C. Wilcox, "Debye potentials," *J. Appl. Math. Mech.*, vol. 6, pp. 167–201, 1957.

[6] C. J. Bouwkamp and H. B. G. Casimir, "On multipole expansions in the theory of electromagnetic radiation", *Physica*, vol. 20, pp. 539–554, 1954.

[7] C. Müller, "Boundary values and diffraction problems", *Symp. Math.* vol. 18, pp. 354–367, 1976.

[8] E. M. Kennaugh, "Multipole Field Expansions and Their Use in Approximate Solutions of Electromagnetic Scattering Problems, Ohio State University, Department of Electrical Engineering, Ph.D. Dissertation, 1959. (Also Report 827–5, Antenna Lab. Ohio State University, Nov. 1959.)

[9] D. R. Rhodes, *Synthesis of Planar Antenna Sources*, Oxford, England: Clarendon Press, 1974.

[10] C. L. Dolph, "A current distribution for broadside arrays which optimizes the relationship between beam width and side lobe level", *Proc. IRE* vol. 34, pp. 335–348, 1946.

[11] R. L. Fante and J. T. Mayhan, "Bounds on the electric field outside a radiating system", *IEEE-AP* vol. 16, pp. 712–717, 1968.

[12] A. Kirsch and P. Wilde, "The optimization of directivity and signal-to-noise ratio of an arbitrary antenna array", *Math. Meth. Appl. Sci.* vol. 10, pp. 153–164, 1988.

[13] A. Kirsch and P. Wilde, "Some remarks concerning a non-quadratic antenna problem", to appear.

[14] T. S. Angell and R. E. Kleinman, "Generalized exterior boundary value problems and optimization for the Helmholtz equation", *J. Optimization Theory Appl.* vol. 37, pp. 469–497, 1982.

[15] T. S. Angell and R. E. Kleinman, "A Galerkin procedure for optimization in radiation problems", *SIAM J. Appl. Math.* vol. 44, pp. 1246–1257, 1984.

[16] T. S. Angell and R. E. Kleinman, "A new optimization method for antenna design", *Ann. des Telecommunications* vol. 40, pp. 341–349, 1985.
[17] S. A. Fast, *An Optimization Method for Solving a Radiation Direction Problem*, University of Delaware, Department of Mathematical Sciences, Ph.D. dissertation, Newark, Delaware, 1988.
[18] H. Burdinsky, *Ein Kontrollproblem bei Ausstrahlungen und Streuungen —Galerkin und bedingtes Gradientenverfahren im Vergleich*, Diploma Thesis, Universität Saarbrücken, West Germany, 1983.

**Thomas S. Angell** was born in California in 1942. He received the B.A. degree in chemistry from Harvard College, Cambride, MA, in 1963 and the M.A. and Ph.D. degrees in mathematics from the University of Michigan, Ann Arbor, in 1964 and 1969, respectively.

In 1969 he joined the faculty at the University of Delaware, Newark where he is currently a Professor of Mathematics. He has held visiting positions at Georgia Institute of Technology, the University of Göttingen, and the University of New Mexico. His main interests are in the theory of optimal control, functional differential equations, integral equations, and the integral equation methods in problems of acoustic and electromagnetic scattering theory.

Dr. Angell is the Associate Editor of the *Journal of Optimization Theory and Applications* and of the *Journal of Mathematical Analysis and Applications*.

**Andreas Kirsch** received the Ph.D. degree in 1978 and the habilitation degree in 1984 from the University of Goettingen, Germany.

From 1986 to 1988 he was with the Institute of Geophysics at Goettingen. Since 1988 he has been a Professor for Applied Mathematics at the University of Erlangen-Nurenberg, Germany. His main areas of research are inverse problems in acoustics and electromagnetic theory.

**Ralph E. Kleinman** was born in New York in 1929. He received the B.A. degree in mathematics from New York University, New York, in 1950, the M.A. degree in mathematics from the University of Michigan, Ann Arbor in 1951, and the Ph.D. degree in mathematics from the Delft University of Technology, The Netherlands, in 1961.

From 1951 to 1968 he was a permanent member of the Radiation Laboratory at the University. He then joined the Department of Mathematical Sciences at the University of Deleware, Newark, where he currently is a Professor and Director of the Center for the Mathematics of Waves. He has held visiting positions at the Technical University of Denmark, University of Strathclyde, Delft University of Technology, AFCRL (now Rome Laboratory), University of Gothenburg, David Taylor Naval Ship Research and Development Center, and the Naval Research Laboratory. His main interests have been concerned with the mathematical problems associated with propagation and scattering of acoustic and electromagnetic waves including radar cross section analysis, integral representations of solutions of the Helmholtz and Maxwell equations, and low frequency perturbation techniques, inverse scattering and iterative solutions of integral equations.

Dr. Kleinman serves as an Associate Editor of *Radio Science* and *Inverse Problems* and is the Chairman, U.S. Commission B of the International Union of Radio Science.

*IMA Journal of Applied Mathematics* (1991) **47**, 265–281

# A Constructive Method for Shape Optimization: A Problem in Hydromechanics

THOMAS S. ANGELL AND RALPH E. KLEINMAN

*Center for the Mathematics of Waves, Department of Mathematical Sciences, University of Delaware, Newark, Delaware 19711, USA*

[Received 12 February 1990 and in revised form 12 August 1990]

Certain hydromechanical quantities associated with a floating or a totally immersed body depend explicitly on the body's geometry. In this paper, the authors consider the problem of choosing the *shape* of the body so that one such quantity, added mass, is optimized. In particular, a constructive method of penalization type is proposed which depends on the availability of a complete family of solutions of the original boundary value problem and it is shown how such families may be generated.

## 1. Introduction

WHEN a body, floating or submerged in an infinite, ideal, inviscid, and irrotational fluid is subjected to a periodic vertical displacement, a wave pattern is created in the fluid. The problem of determining this pattern from a knowledge of the body geometry and applied forces is well known in fluid mechanics.

In problems with either partly or fully submerged objects, quantities of physical interest are not only the wave patterns which may be derived from the velocity potential but also functionals of the potential such as added mass and damping factors which measure the distribution of energy in the fluid (see e.g. Wehausen & Laitone, 1960: p. 567). These factors are dependent on the body geometry and the natural question arises as to whether such quantities may be optimized over restricted classes of body geometry.

The question of optimizing the added mass or similar functionals by choosing the shape of the object was addressed by Angell *et al.* (1986), who established the existence of an optimal shape for a totally submerged body for a fluid of finite depth in an appropriate function-space setting. This problem is again considered in the present paper, this time presenting a constructive method for actually finding shapes which optimize added mass or damping. In the terminology of optimal control, the problem is one of optimization of geometrical elements (see e.g. Lions, 1972). Other optimization problems of this general class have been studied previously by, for example, Cea and co-workers (1974, 1975), Chenais (1975), and Pironneau (1973, 1974). However, in contrast to much of this earlier work, the natural setting for our problem is in an unbounded rather than in a bounded domain.

It will come as no surprise to those familiar with the peculiar difficulties associated with exterior boundary value problems that it is particularly useful to reformulate the original problem, which here includes not only boundary

conditions given on the bounded surface of the body but also those on the free surface and on the bottom (both of which are of infinite extent), as a uniquely solvable integral equation defined on the boundary of the body. The efficacy of the boundary integral equation approach depends, at least in the first instance, on the uniqueness of solutions to the original boundary value problem. This uniqueness question should not be confused with the question of unique solvability of some boundary integral equations derived, say, from a layer ansatz. This latter question is sometimes referred to as the problem of irregular frequencies.

The unique solvability of the boundary value problem for the floating body is not completely understood, although, as John remarked in his fundamental paper (John, 1950: p. 49), 'There appears to be no physical reason why . . . the primary wave motion together with the motion of the obstacle should not determine the motion in the liquid uniquely'. In that paper, John established uniqueness only with certain restrictions on the body shape, in particular that it be convex and smooth, that it have normal intersection with the free surface, and moreover that vertical rays from the free surface intersect the body at most once. These conditions may be relaxed somewhat (Kleinman, 1982; Simon & Ursell, 1984) but some nonphysical restrictions remain.

When the body is completely submerged, John's uniqueness proof no longer applies. However, Maz'ja (1978) has provided a proof for a class of bodies delimited once again by certain geometric restrictions. The recent and interesting paper of Hulme (1984) discusses the result of Maz'ja and effectively describes the geometric meaning of the result. We will give a precise statement of this result in the next section. At this point, suffice it to say that Maz'ja's condition provides a reasonable class of bodies for which we can assert the uniqueness of solutions of the boundary value problem in the case when the body is totally submerged.

In the case of the totally submerged body, Angell *et al.* (1986) derived, by using a Green's function, an integral equation which is uniquely solvable for all frequencies. This Green's function, introduced by John, is that appropriate to the entire fluid domain with no body present and satisfying the boundary conditions at the bottom of the fluid (assumed flat) and the linearized free-surface condition on the entire fluid/air boundary. It is the formulation of the boundary value problem, the statement of Maz'ja's theorem, and the derivation of this boundary integral equation that are summarized in the next section, while Section 3 contains a description of the optimization problem and a statement of the results obtained in Angell *et al.* (1986) concerning the existence of an optimal body shape.

It is the final section, Section 4, which contains our main results. There we turn to the question of a constructive method for finding approximate optimal surfaces. We prove that certain families of functions form complete families of solutions and propose a penalization-type method for the constructive solution. The idea of using complete families to find approximate solutions to elliptic equations goes back at least to the work of Picone and of Fichera (see Miranda, 1970, for references). Angell & Kleinman (1984, 1985) have used such families in treating some optimization problems which arise in acoustic and in electromagnetic radiation problems. An approximation method similar to that proposed here

is discussed in the context of an inverse transmission problem by Angell *et al.* (1987). A related method in the inverse acoustic problem has been reported by Kirsch & Kress (1986).

## 2. The exterior boundary value problem

We are concerned with solutions of Laplace's equation in a domain $D \subset \mathbb{R}^3$, unbounded in the $x$ and $z$ directions and exterior to a bounded boundary $\Gamma$, which is assumed to be a Lyapunov surface of index 1. A Cartesian coordinate system is fixed with the origin on the free surface and in terms of which the domain $D^+ = (\mathbb{R}^2 \times [-h, 0]) \backslash (\Gamma \cup D^-)$, where $D^-$ denotes the interior of the submerged body, as indicated in Fig. 1.

The submerged body will be assumed to be simply connected and lie in a strip $\mathbb{R}^2 \times [-h + \epsilon_0, -\epsilon_0]$, with $\epsilon_0 > 0$. The condition that the surface be Lyapunov of index 1 guarantees, among other things, that there exists a Lipschitz continuous normal $\hat{n}$ at all points of $\Gamma$. We emphasize that $\hat{n}$ is oriented so that it points into $D^+$. Points will be denoted by $p = (x_p, y_p, z_p)$ and, in cylindrical coordinates, by $p = (\rho_p, \theta_p, y_p)$. And the subscripts will be omitted if there is no danger of confusion.

With these conventions in mind, we consider the boundary value problem

$$\Delta \phi = 0 \qquad \text{in } D^+, \tag{2.1a}$$

$$\frac{\partial \phi}{\partial n} + k\phi = 0 \quad \text{on } y = 0, \tag{2.1b}$$

$$\frac{\partial \phi}{\partial n} = 0 \qquad \text{on } y = -h, \tag{2.1c}$$

$$\frac{\partial \phi}{\partial n} = G \qquad \text{on } \Gamma, \tag{2.1d}$$

together with a radiation condition

$$\frac{\partial \phi}{\partial \rho} - ik_0\phi = O(\rho^{-\frac{1}{2}}). \tag{2.1e}$$

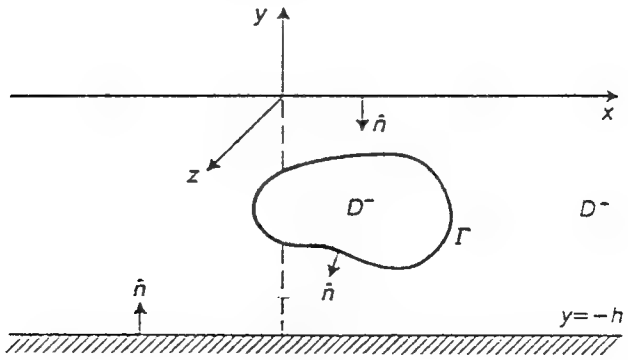



FIG. 1.

In this formulation, $G \in C(\Gamma)$ and $k := \omega^2/g$ is a real parameter, where $\omega$ is the frequency of an oscillation (assumed time periodic) and $g$ is the gravitational constant, while $k_0$ is the root with the largest real part of the transcendental equation

$$k_n \sinh k_n h = k \cosh k_n h. \tag{2.2}$$

Maz'ja (1978) introduced a restricted class of boundaries for which this boundary value problem has at most one solution. We formulate that theorem as follows.

THEOREM 2.1 *Let $V$ be the vector field in $\mathbb{R}^3$ defined by*

$$V = \frac{\rho(y^2 - \rho^2)}{\rho^2 + y^2}\hat{\rho} - \frac{2\rho^2 y}{\rho^2 + y^2}\hat{y}.$$

*Then the homogeneous boundary value problem* (2.1) *with $G = 0$ has only the trivial solution provided that*

$$V \cdot \hat{n} \geqslant 0 \quad \text{on } \Gamma. \tag{2.3}$$

A discussion of this result and its geometric significance may be found in Hulme (1984). We will refer to the class of all such surfaces as the Maz'ja class.

Following John (1950), we introduce the Green's function for this problem, which is normalized to have the form

$$\gamma(p, q) = -\frac{1}{2\pi}\frac{1}{|p - q|} + R(p, q), \tag{2.4}$$

where the function $R$ has bounded derivatives with respect to $q$ for points $p \in \Gamma$ (see John, 1950: p. 96) and $\gamma$ satisfies conditions (2.1b, c, e). Using this Green's function to define single and double layer potentials, the usual jump conditions can be established as in the potential-theoretic case since the singular behaviour of $\gamma$ and $\partial\gamma/\partial n_q$ is determined by the first term in (2.4). For convenience, we record these results here:

$$\lim_{p \to \Gamma^{\pm}} \frac{\partial}{\partial n_p} \int_\Gamma u(q)\gamma(p, q)\, d\Gamma_q = \pm u(p) + \int_\Gamma u(p) \frac{\partial \gamma(p, q)}{\partial n_p}\, d\Gamma_q, \tag{2.5}$$

$$\lim_{p \to \Gamma^{\pm}} \int_\Gamma u(q) \frac{\partial}{\partial n_q}\gamma(p, q)\, d\Gamma_q = \mp u(p) + \int_\Gamma u(p) \frac{\partial \gamma(p, q)}{\partial n_q}\, d\Gamma_q, \tag{2.6}$$

where $p \to \Gamma^{\pm}$ means $p$ approaches $\Gamma$ from $D^+$ or from $D^-$, $u \in L^2(\Gamma)$, and the relations (2.5) and (2.6) hold in the $L^2$ sense (Miranda, 1970).

Moreover, if $\phi$ is a solution of the boundary value problem (2.1), then one may use Green's theorem to establish the familiar relation

$$\int_\Gamma \left(\gamma(p, q)\frac{\partial}{\partial n_q}\phi(q) - \phi(q)\frac{\partial}{\partial n_q}\gamma(p, q)\right) d\Gamma_q = \begin{cases} 2\phi(p) & (p \in D^+) \\ \phi(p) & (p \in \Gamma), \\ 0 & (p \in D^-). \end{cases} \tag{2.7}$$

If one then uses the boundary condition (2.1d), we have, for $p \in \Gamma$,

$$\int_\Gamma \gamma(p, q)G(q)\,d\Gamma_q - \int_\Gamma \phi(q)\frac{\partial}{\partial n_q}\gamma(p, q)\,d\Gamma_q = \phi(p), \tag{2.8}$$

or, in operator notation,

$$(I + \bar{K}^*)\phi = \int_\Gamma \gamma(p, q)G(q)\,d\Gamma_q, \tag{2.9}$$

where $\bar{K}^*$ is the boundary integral operator with kernel $\partial\gamma/\partial n_q$. We pause to remark that, given a solution $u$ of this integral equation, we may represent the solution of the boundary value problem according to the relation (2.7) by

$$\phi(p) = \tfrac{1}{2}\int_\Gamma \gamma(p, q)G(q)\,d\Gamma_q - \tfrac{1}{2}\int_\Gamma u(q)\frac{\partial}{\partial n_q}\gamma(p, q)\,d\Gamma_q \quad (p \in D^-), \tag{2.10}$$

and, again using the jump relations, one sees easily that

$$\phi|_\Gamma = u, \tag{2.11}$$

which is a direct relationship between the solution of the boundary integral equation and the boundary values taken on by the solution. As we will see below when we consider the optimization problem, it is particularly convenient to have this formulation since the cost functional involves just the trace of the solution of $\phi$ of (2.1) on $\Gamma$. Such a direct relation does not obtain when one uses a layer approach in which one assumes that the solution $\phi$ has a representation as a single layer,

$$\phi(p) = \int_\Gamma u(q)\gamma(p, q)\,d\Gamma_q,$$

and then uses the boundary condition and jump relations to obtain an integral equation for $u$.

With the aid of these jump conditions, we have proved the unique solvability of the boundary integral equation (2.9). Specifically, we may state the following theorem, referring to Angell *et al.* (1986) for the proof.

THEOREM 2.2 *Let $\Gamma$ be Lyapunov of index 1 and belong to the Maz'ja class. Let $G \in C(\Gamma)$. Then the integral equation* (2.9) *has a unique solution in* $L^2(\Gamma)$.

*Remark.* In fact, using a standard argument, the solution whose existence is guaranteed by this last theorem can be shown to be continuous since $G \in C(\Gamma)$.

### 3. The optimization problem

Let $\Gamma_0 = \{p \in \mathbb{R}^3 : |p| = 1\}$ denote the surface of the unit ball in $\mathbb{R}^3$ and let $C^{1,1}(\Gamma_0)$ denote the space of continuously differentiable functions whose first derivatives satisfy a Lipschitz condition and which is equipped with the usual Hölder norm $\|\cdot\|_{1,1}$ (see e.g. Colton & Kress, 1983). We will assume that we are

given a family of surfaces which can be described by $C^{1,1}$ parametrizations,

$$\Gamma(f) = \left\{ p \in \mathbb{R}^3 : p = f(\bar{p})\bar{p} + p_0, \bar{p} = \frac{p - p_0}{|p - p_0|} \right\}, \tag{3.1}$$

where $f : \Gamma_0 \to \mathbb{R}^3$ is an element of $C^{1,1}(\Gamma_0)$ and $p_0 \in \mathbb{R}^2 \times (-h + \epsilon_0, -\epsilon_0)$. Let $a$ and $b$ be two positive constants and define the subset $\mathcal{F}_{a,b} \subset C^{1,1}(\Gamma_0)$ by

$$\mathcal{F}_{a,b} = \{ f \in C^{1,1}(\Gamma_0):$$

$$\|f\|_{1,1} \leq b, f(\bar{p})\bar{p} + p_0 \in \mathbb{R}^2 \times (-h + \epsilon_0, -\epsilon_0), \quad f(\bar{p}) \geq a \ (\bar{p} \in \Gamma_0) \}. \tag{3.2}$$

DEFINITION 3.1 *A surface $S$ in $\mathbb{R}^3$ will be called* admissible *provided $S$ can be described by a parametrization $f \in \mathcal{F}_{a,b}$ and $S$ is contained in the Maz'ja class (cf. Theorem* 1.1).

Note that, since each admissible surface is completely determined by the function $f$, we will henceforth simply refer to 'the surface $f$', although, when convenient, we will use the notation $\Gamma(f)$. Clearly, each admissible surface describes a surface bounding a bounded region which contains a ball of radius $\frac{1}{2}a$ and centre $p_0$ in its interior (see Fig. 2). We will, when necessary, denote the region in the domain $\mathbb{R}^2 \times (-h, 0)$ exterior to an admissible surface $f$ by $D_f^-$ and the interior of the surface by $D_f^+$.

Now we limit attention to a compact subset $U_{ad}$ of the class of admissible surfaces. Since the embedding $C^2(\Gamma_0) \to C^{1,1}(\Gamma_0)$ is compact, we may choose $U_{ad}$ to be a closed subset of functions in $C^2(\Gamma_0)$. This particular choice leads to a nonlinear optimization problem over a closed convex set. The convexity will be advantageous for subsequent numerical considerations. We note, however, that the subsequent results could be achieved using any compact subset of the class of admissible functions.

We now wish to consider a *family* of boundary value problems of the type

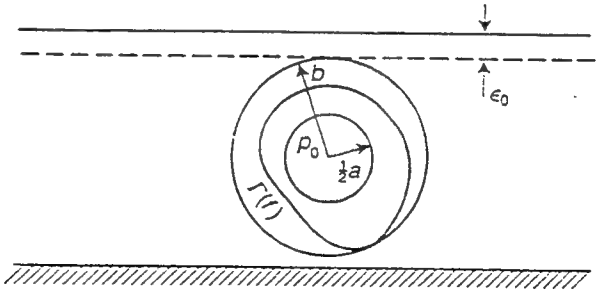



FIG. 2.

discussed in Section 2 which may be considered as indexed by $U_{ad}$:

$$\Delta\phi(p) = 0 \quad (p \in D_f^+), \tag{3.3a}$$

$$\frac{\partial\phi}{\partial n} + k\phi = 0 \quad \text{on } y = 0, \tag{3.3b}$$

$$\frac{\partial\phi}{\partial n} = 0 \quad \text{on } y = -h, \tag{3.3c}$$

$$\frac{\partial\phi}{\partial n} = G \quad \text{on } \Gamma(\mathrm{f}), \tag{3.3d}$$

$$\frac{\partial\phi}{\partial\rho} - ik_0\phi = o(\rho^{-\frac{1}{2}}). \tag{3.3e}$$

Note that, because we are considering a family of boundary value problems, the data $G$ in (3.3d) must be defined throughout the domain formed by the union of all admissible surfaces. This is indeed the case for heaving motion, where $G = -\hat{n}\cdot\hat{y}$.

With this understanding, each choice of surface $f \in U_{ad}$ gives rise, according to Theorem 2.1, to a potential $\phi = \phi(p;f)$, with $p \in D_f^+$. Denote the trace of $\phi(p;f)$ on $f$ by

$$\bar{\phi}_f(\bar{p}) := \phi\big(f(\bar{p})\bar{p} + p_0; f\big) \quad (\bar{p} \in \Gamma_0). \tag{3.4}$$

The class of optimization problems that we discuss below involves a functional defined as follows. Let $L: C(\Gamma_0) \to \mathbb{R}$ be continuous and note that, since $\bar{\phi}_f(\bar{p}) \in C(\Gamma_0)$, this functional may also be considered as a mapping from $U_{ad}$ into $\mathbb{R}$ by restricting the domain of $L$ to $\{\bar{\phi}_f \in C(\Gamma_0) : f \in U_{ad}\}$. In this sense, we define

$$L[f] := L(\bar{\phi}_f) : U_{ad} \to \mathbb{R}. \tag{3.5}$$

We seek a function $f_0 \in U_{ad}$ such that

$$L[f_0] \leqslant L[f] \quad \text{for all } f \in U_{ad} \tag{3.6}$$

or

$$L[f_0] \geqslant L[f] \quad \text{for all } f \in U_{ad}. \tag{3.7}$$

We will confine our discussion to the problem of minimization. This is sufficiently general since the problem of maximizing a functional $L$ may always be replaced by that of minimizing $-L$.

Specific forms of the functional $L$ of (3.5) may be chosen to reflect desirable design criteria. For example, as mentioned in the introduction, one may choose $L$ to represent the added mass of the hull. In this case, the problem of interest is that of minimizing the functional $L$ in order to reduce the hydrodynamic force on the ship hull, a goal of obvious importance to ship design. Indeed, it is well known (see Wehausen & Laitone, 1960: pp. 563–7) that the added mass of a

particular hull may be represented by

$$M_a = \mathrm{Re} \int_{\Gamma(f)} \phi(p) \frac{\partial}{\partial n_p} \phi(p) \, d\Gamma_p,$$

where Re stands for the real part of the integral. This form, in light of the boundary condition (3.3d), leads to the functional

$$L[f] = \mathrm{Re} \int_{\Gamma_0} \bar{\phi}(p; f) G(f(\bar{p})\bar{p} + p_0) J_f(\bar{p}) \, d\Gamma_0, \tag{3.8}$$

where $J_f$ is the Jacobian of the transformation $p = f(\bar{p})\bar{p} + p_0$, and $d\Gamma_0 = d\theta \, d\varphi$, with $\theta$ and $\phi$ the spherical polar angles of the point $\bar{p}$. We remark that, if we can show that the map $f \to \bar{\phi}(\cdot; f)$ is continuous as a mapping from $U_{ad}$ to $C(\Gamma_0)$, then, regardless of the particular form of the functional $L$, its continuity together with the assumed compactness of $U_{ad}$ will guarantee the existence of an optimal solution.

More generally, we may consider the functional

$$L = \sum_{i,j=1}^{7} u_i u_j m_{ij}, \tag{3.9}$$

where

$$m_{ij} := \rho \int_\Gamma \phi_j \frac{\partial}{\partial n} \phi_i \, d\Gamma \quad (i,j = 1,\ldots,7) \tag{3.10}$$

are the components of the added mass tensor, the $u_i$ $(i = 1,\ldots,7)$ represent the velocity components (assumed given) of the body, $\rho$ is the density of the fluid, and each $\phi_i$ represents a velocity potential of a rigid-body motion with unit amplitude in the absence of incident waves (see Newman, 1977: pp. 287–8). We remark that, while this functional is quadratic in the $u_i$, it is *not* quadratic when considered as a functional of the surface $\Gamma$. Introduction of this functional permits optimization with respect to combinations of the added masses, perhaps omitting some, but does not change the analysis below, since it is trivial to rewrite the functional (3.9) in the form

$$L = \rho \int_\Gamma \phi \frac{\partial}{\partial n} \phi \, d\Gamma \tag{3.11}$$

for $\phi = \sum_{i=1}^{7} u_i \phi_i$, the harmonic function $\phi$ satisfying all required boundary conditions.

Angell *et al.* (1986) proved the continuity of the mapping $f \to \bar{\phi}(\cdot; f)$ from $U_{ad}$ to $C(\Gamma_0)$. By introducing the functions

$$\bar{\phi}(\bar{p}) := \phi(f(\bar{p})\bar{p} + p_0), \tag{3.12}$$

$$G_f(\bar{p}) := G(f(\bar{p})\bar{p} + p_0), \tag{3.13}$$

and kernels

$$a_f(\bar{p}, \bar{q}) := \frac{\partial}{\partial n_q} \gamma(f(\bar{p})\bar{p} + p_0, f(\bar{q})\bar{q} + p_0) J_f(\bar{q}), \tag{3.14}$$

$$b_f(\bar{p}, \bar{q}) := \gamma(f(\bar{p})\bar{p} + p_0, f(\bar{q})\bar{q} + p_0) J_f(\bar{q}), \tag{3.15}$$

the integral equation (2.9) may be rewritten as

$$\bar{\phi}(\bar{p}) + \int_{\Gamma_0} a_f(\bar{p}, \bar{q})\bar{\phi}(\bar{q})\, d\Gamma_{\bar{q}} = \int_{\Gamma_0} b_f(\bar{p}, \bar{q}) G_f(\bar{q})\, d\Gamma_{\bar{q}}. \tag{3.16}$$

The integral operators $A_f$ and $B_f$ defined by the kernels $a_f$ and $b_f$ are compact operators on $C(\Gamma_0)$, a fact established in Angell *et al.* (1986).

The basic results are the following theorem and its corollary which we will use in the next section.

THEOREM 3.1 *Let* $B(C(\Gamma_0))$ *denote the space of bounded linear operators on* $C(\Gamma_0)$ *equipped with the uniform operator topology and assume that the map* $f \to G_f$ *from* $C^{1,1}(\Gamma_0)$ *into* $C(\Gamma_0)$ *is Hölder continuous. Then the mappings* $f \to A_f$ *and* $f \to B_f$ *of* $C^{1,1}(\Gamma_0)$ *into* $B(C(\Gamma_0))$ *are Hölder continuous. Moreover, since the set* $U_{ad}$ *is compact, the map* $f \to \bar{\phi}(\cdot; f)$ *of* $U_{ad}$ *into* $C(\Gamma_0)$, *where* $\bar{\phi}(\cdot; f)$ *is the unique solution of the boundary integral equation* (3.16), *is Hölder continuous.*

The continuity of the mappings $f \to \bar{\phi}(\cdot; f)$ and $f \to G_f$ lead immediately to the result that the optimization problem defined by (3.6) has a solution. For the particular case of the added mass functional (3.8), which is the functional we will concentrate on in Section 4, we may state the following corollary to Theorem 3.1.

COROLLARY 3.1 *Under the hypotheses of Theorem* 3.1, *the functional* $L[\cdot]$ *defined by the equation* (3.8) *is continuous as a map from* $U_{ad}$ *into* $\mathbb{R}$ *and consequently takes on its absolute minimum on the set* $U_{ad}$.

## 4. A penalization method

With the groundwork in place, we turn to the main results, the development of a constructive procedure, a penalization method, for finding approximate solutions of the optimization problem described in Section 3. For the sake of definiteness, we will formulate the procedure in terms of the specific functional (3.8). It will become clear that the method is applicable to a wide class of functionals of which (3.8) is but one example. Such methods have been applied by others to systems governed by partial differential equations (see e.g. Lions 1971, 1972). Generally, they involve the introduction of additional terms to the cost functional involving both the partial differential operator and the various initial and boundary conditions. Here, we propose to carry out the minimization, not over an entire Sobolev space as in earlier applications, but over a compact set of functions whose traces on the class of admissible surfaces serve as boundary data for exterior solutions of the boundary value problem. With this approach, we will need to introduce only one penalization term corresponding to (3.3d).

We will then turn to the development of a Galerkin-type procedure based on the use of complete families of solutions. Elements of such a family are harmonic functions, defined in the region $(\mathbb{R}^2 \times [-h, 0]) \backslash B_{a_0}$, where $B_{a_0}$ is a ball of radius $a_0 < \frac{1}{2}a$, satisfying not only the boundary conditions (3.3b) and (3.3c) on the free surface and bottom respectively but also the radiation condition (3.3e). The use of complete families not only simplifies the form of the cost functional but also

offers the significant advantage of allowing us to avoid the difficulty, common in numerical procedures for inverse problems, of having to solve a succession of direct problems as an essential step in an iterative procedure. The method given here, once the dimension of the approximating subspace is fixed, produces both a suboptimal surface and the appropriate field exterior to that surface.

We will proceed in three steps. First, we will set up the penalized problem in the infinite-dimensional setting. Second, we will study the finite-dimensional problems generated by considering subspaces spanned by finite collections of a complete family of solutions, and proving convergence of minimizers of the finite-dimensional problems to a solution of the original problem. Finally, we will show how such a complete family may be obtained so that the procedure may be implemented.

Let $A$ denote the closed annular domain $\overline{B_b \backslash B_a}$ lying in the strip $\mathbb{R}^2 \times (-h, 0)$ which is determined by the spheres $B_b$ and $B_a$, where $a$ and $b$ are the constants appearing in the definition of the class $\mathcal{F}_{a,b}$. Thus all the admissible surfaces $\Gamma_f$ lie in $A$. We will assume, in concert with the remarks following equations (3.3), that there exists a function $H \in C^2(A_\eta)$ (i.e. $C^2$ in an $\eta$-neighbourhood of $A$) such that $\partial H/\partial n = G$ on $\Gamma_f$ for each $f \in U_{ad}$.

Let $M$ be any constant satisfying $\|H\|_{C^2(A_\eta)} \leq M$ and define

$$S_M = \{F \in C^2(A) : \|F\|_{C^2(A_\eta)} \leq M\}, \tag{4.1}$$

where $\|\cdot\|_{C^2(A_\eta)}$ is the usual $C^2$-norm. Since the embedding $C^2(A) \to C^1(A)$ is compact, $S_M$ is relatively compact in $C^1(A)$. If we denote its $C^1$-closure by $\hat{S}$, then $\hat{S}$ is compact in $C^1(A)$.

For every $F \in \hat{S}$ and $f \in U_{ad}$, we can consider the boundary value problem with boundary data

$$\frac{\partial F}{\partial n} \quad \text{on } \Gamma_f. \tag{4.2}$$

By Theorem 2.2, the corresponding integral equation (2.9) with $G$ replaced by $\partial F/\partial n$ has a unique solution $u_{f,F}$, which can be used to generate a unique solution $\phi_{f,F}$ to the boundary value problem by using (2.10). This is true, in particular, for $F = H$.

We now introduce three functionals $l_1$, $l_2$, and $L_\nu$ defined on the compact set $U_{ad} \times \hat{S}$ by

$$l_1(f, F) := \operatorname{Re} \int_{\Gamma_f} \phi_{f,F}(p) \frac{\partial}{\partial n} F(p)\, d\Gamma \tag{4.3}$$

and

$$l_2(f, F) := \left\| \frac{\partial F}{\partial n} - G \right\|^2_{L^2(\Gamma_f)}, \tag{4.4}$$

while, for a given $\nu > 0$,

$$L_\nu(f, F) := l_1(f, F) + \nu l_2(f, F). \tag{4.5}$$

With the usual reparametrization, this expression can be rewritten in terms of integrals over $\Gamma_0$:

$$L_\nu(f, F) = \mathrm{Re} \int_{\Gamma_0} \phi_{f,F}(p) \frac{\partial}{\partial n} F(p)\Big|_{p=p_0+f(\hat{q})\hat{q}} J_f(\hat{q})\, d\Gamma_0 + \nu \int_{\Gamma_0} \left| \frac{\partial}{\partial n} F(p) - G(p) \right|^2_{p=p_0-f(\hat{q})\hat{q}} J_f(\hat{q})\, d\Gamma_0. \quad (4.6)$$

The functional $l_2$ is intended to ensure that $F$ is chosen to approximate the given data, while $l_1$ ensures that the added mass is minimized. Indeed $l_2$ can be viewed as a penalty term which penalizes deviations from the desired boundary condition on the surface $f$. Using arguments completely analogous to those used to prove Theorem 3.1, it can be shown that both $l_1$ and $l_2$ are continuous on $U_{ad} \times \hat{S}$, and hence there exists a pair $(f_\nu, F_\nu) \in U_{ad} \times \hat{S}$ such that

$$L_\nu(f_\nu, F_\nu) \leq L_\nu(f, F) \quad \text{for all } (f, F) \in U_{ad} \times \hat{S}.$$

We note that, if we consider an increasing sequence of penalization parameters $\{\nu_m\}$ such that $\nu_m \to \infty$ as $m \to \infty$, the corresponding sequence of optimal solutions will contain at least a subsequence $\{(f_k, F_k)\}$ which will converge to an element $(f_0, F_0) \in U_{ad} \times \hat{S}$. A standard argument (see e.g. Luenberger, 1969: p. 305) shows that, in fact, $(f_0, F_0)$ is a minimizer for the original optimization problem (3.6), so that $\partial F_0/\partial n = G$.

As it stands, the functional (4.6) suffers from the drawback that there is no way to associate the added mass with a particular surface without first solving the direct problem for $\phi_{f,F}$. We now propose a Galerkin-type approximation method which eliminates the need for first solving the direct problem. In fact, if the dimension of the approximating subspace is fixed, then the approximate solution of the minimization problem is obtained by simultaneously solving for $F$ and the optimal surface without requiring the solution of a succession of direct problems.

The approximating subspaces will be defined in terms of a countable family $\mathcal{H}$ of harmonic functions defined as follows. The elements $v_l$ of $\mathcal{H}$ are harmonic in $(\mathbb{R}^2 \times [-h, 0]) \backslash B_{a_0}$ and satisfy conditions (3.3b), (3.3c), and (3.3e), and the normal derivatives $\{\partial v_l/\partial n|_{\Gamma(f)}\}_{l=1}^{\infty}$ are linearly independent and are complete in $L^2(\Gamma(f))$ for all $f \in U_{ad}$. We will show one way to construct such a family at the end of this section. Postponing that analysis, we begin the description of the approximation procedure by establishing a convergence result which we will use in the proof of our main result. We remark that the completeness of the family $\mathcal{H}$ allows us to approximate any function in $L^2(\Gamma(f))$ as closely as desired with a finite linear combination of the normal derivatives $\partial v_l/\partial n$. Even more is true as a consequence of this choice of the functions $v_l$ as harmonic functions. This additional approximation result is described in the next statement in which $\Gamma$ is a fixed surface.

LEMMA 4.1 *Let $G \in L^2(\Gamma)$ and let $\phi$ be a solution of the system* (3.3) *with $\partial\phi/\partial n = G$. Let $\{v_l\}_{l=1}^{\infty}$ be the complete family of harmonic functions as described above. Suppose that, for each integer $N \geq 1$, the coefficients $\{c_l^{(N)} : l = 1, 2, \ldots, N\}$*

*are chosen so as to minimize* $\|\sum_{l=1}^{N} c_l^{(N)}(\partial v_l/\partial n) - G\|_{L^2(\Gamma)}$. *Then*

$$\lim_{N\to\infty} \left\| \sum_{l=1}^{N} c_l^{(N)} v_l - \phi \right\|_{L^2(\Gamma)} = 0, \tag{4.7}$$

*where* $\phi$ *is the solution of the integral equation* (2.9).

*Proof.* Since $\phi$ and $v_l$ $(l = 1, 2, \ldots)$ are harmonic and satisfy the radiation condition, Green's theorem yields the relations

$$(I + \bar{K}_\Gamma^*)v_l = \int_\Gamma \gamma(p, q) \frac{\partial v_l}{\partial n}(q)\, d\Gamma_q, \tag{4.8}$$

$$(I + \bar{K}_\Gamma^*)\phi = \int_\Gamma \gamma(p, q) G(q)\, d\Gamma_q, \tag{4.9}$$

$\bar{K}_\Gamma^*$ being the double layer operator associated with the surface $\Gamma$ (see (2.7)–(2.9)). We may conclude immediately that, since the integral equation (2.9) has a unique solution,

$$\sum_{l=1}^{N} c_l^{(N)} v_l - \phi = (I + \bar{K}_\Gamma^*)^{-1} \int_\Gamma \gamma(p, q) \left( \sum_{l=1}^{N} c_l^{(N)} \frac{\partial v_l}{\partial n}(q) - G(q) \right) d\Gamma_q. \tag{4.10}$$

Moreover,

$$\left\| \sum_{l=1}^{N} c_l^{(N)} v_l - \phi \right\|_{L^2(\Gamma)} \leq \|(I + \bar{K}_\Gamma^*)^{-1}\|_{L^2(\Gamma)} \|S_\Gamma\| \left\| \sum_{l=1}^{N} c_l^{(N)} \frac{\partial v_l}{\partial n} - G \right\|_{L^2(\Gamma)}, \tag{4.11}$$

where $S_\Gamma$ is the single layer operator associated with $\Gamma$. The result follows from the boundedness of the two operators and the completeness of the $\partial v_l/\partial n$.

The elements of the family $\mathcal{H}$ individually satisfy the free-surface condition, the radiation condition, and the boundary condition on the sea floor, while their normal derivatives are complete and linearly independent on $L^2(\Gamma(f))$ for all $f \in U_{ad}$. This makes them useful, not only in approximating solutions of the submerged body problem (3.3), but also in formulating a sequence of approximate optimization problems.

To this end, it is convenient to introduce the subspace $V^N \subset \mathcal{H}$ spanned by the functions $v_l \in \mathcal{H}$ $(1 \leq l \leq N)$. We then consider the set $S^N = V^N \cap S_M$, where $S_M$ is defined in (4.1). In terms of this set, we may define an approximate inverse problem as follows. For a given integer $N$ and function $G$, find $w \in S^N$ and $f \in U_{ad}$ which together minimize

$$L_\nu^{(N)}[f, w] := \mathrm{Re} \int_{\Gamma_0} \phi_{f,w}(p) \frac{\partial}{\partial n} w(p) \Big|_{p=p_0+f(\bar{q})\bar{q}} J_f(\bar{q})\, d\Gamma_0 + \nu \int_{\Gamma_0} \left| \frac{\partial}{\partial n} w(p) - G(p) \right|_{p=p_0+f(\bar{q})\bar{q}} J_f(\bar{q})\, d\Gamma_0.$$

The question immediately arises as to how the solution of the approximate optimization problem is related to an optimal solution of the exact problem. As

the following result shows, the cluster points of optimal solutions of the approximate problems are solutions of the exact optimization problem.

THEOREM 4.1 *Let*

$$L_\nu^0 := \min\{L_\nu[f, F] : (f, F) \in U_{ad} \times \hat{S}\},$$
$$L_\nu^{(N)^0} := \min\{L_\nu^{(N)}[f, \Phi] : (f, \Phi) \in U_{ad} \times S^N\}.$$

*Then, for fixed* $\nu$, $\lim_{N\to\infty} L_\nu^{(N)^0} = L_\nu^0$. *Furthermore, if* $f_\nu^{(N)}$ *is an optimal surface for* $L_\nu^{(N)}$ *with corresponding optimal* $\Phi_\nu^{(N)}$, *then every cluster point of the sequence* $\{f_\nu^{(N)}, \Phi_\nu^{(N)}\}$ *in* $C^{1,\alpha}(\Gamma_0) \times C^1(A)$ *is optimal for* $L_\nu$.

*Proof.* Suppose that $\{\bar{f}_\nu, \bar{\Phi}_\nu\} \in U_{ad} \times \hat{S}$ is optimal for $L_\nu$, so that $L_\nu^0 = L_\nu[\bar{f}_\nu, \bar{\Phi}_\nu]$, and let $(f_\nu^0, \Phi_\nu^0)$ be a cluster point of the sequence $\{(f_\nu^{(N)}, \Phi_\nu^{(N)})\}_{N=1}^\infty$ in $C^{1,\alpha}(\Gamma_0) \times C^1(A)$. Thus there is a subsequence $\{(f_\nu^{(N_k)}, \Phi_\nu^{(N_k)})\}_{k=1}^\infty$ which converges to $(f_\nu^0, \Phi_\nu^0)$. By continuity, we have

$$\lim_{k\to\infty} L_\nu(f_\nu^{(N_k)}, \Phi_0^{(N_k)}) = L_\nu(f_\nu^0, \Phi_\nu^0) \geq L_\nu^0 = L_\nu(\bar{f}_\nu, \bar{\Phi}_\nu).$$

We wish to show that, indeed, $L_\nu(f_\nu^0, \Phi_\nu^0) \leq L_\nu^0$.

Note first that since $(f_\nu^{(N_k)}, \Phi_\nu^{(N_k)})$ is optimal for $L_\nu^{(N_k)}$, we have the estimate

$$L_\nu^{(N_k)}(f_\nu^{(N_k)}, \Phi_\nu^{(N_k)}) \leq L_\nu^{(N_k)}(\bar{f}_\nu, \Psi) \quad \text{for all } \Psi \in S^{N_k}.$$

Hence

$$\begin{aligned}
&L_\nu^{(N_k)}(f_\nu^{(N_k)}, \Phi_\nu^{(N_k)}) \\
&\leq \mathrm{Re}\int_{\Gamma(\bar{f}_\nu)} \phi_{\bar{f}_\nu,\Psi}\frac{\partial\Psi}{\partial n}\,d\Gamma(\bar{f}_\nu) + \nu\left\|\frac{\partial\Psi}{\partial n} - G\right\|^2_{L^2(\Gamma(\bar{f}_\nu))} \\
&\leq \int_{\Gamma(\bar{f}_\nu)} |\phi_{\bar{f}_\nu,\Psi}|\left|\frac{\partial\Psi}{\partial n} - \frac{\partial\bar{\Phi}_\nu}{\partial n}\right| d\Gamma(\bar{f}_\nu) + \int_{\Gamma(\bar{f}_\nu)}\left|\frac{\partial\bar{\Phi}_\nu}{\partial n}\right| |\phi_{\bar{f}_\nu,\Psi} - \phi_{\bar{f}_\nu,\bar{\Phi}_\nu}|\,d\Gamma(\bar{f}_\nu) \\
&\quad + \mathrm{Re}\int_{\Gamma(\bar{f}_\nu)} \phi_{\bar{f}_\nu,\bar{\Phi}_\nu}\frac{\partial\bar{\Phi}_\nu}{\partial n}\,d\Gamma(\bar{f}_\nu) + \nu\left(\left\|\frac{\partial\Psi}{\partial n} - \frac{\partial\bar{\Phi}_\nu}{\partial n}\right\|^2_{L^2(\Gamma(\bar{f}_\nu))} + \left\|\frac{\partial\bar{\Phi}_\nu}{\partial n} - G\right\|^2_{L^2(\Gamma(\bar{f}_\nu))}\right) \\
&\leq \left(\int_{\Gamma(\bar{f}_\nu)} |\phi_{\bar{f}_\nu,\Psi}|^2\,d\Gamma(\bar{f}_\nu)\right)^{\frac{1}{2}}\left(\int_{\Gamma(\bar{f}_\nu)}\left|\frac{\partial\Psi}{\partial n} - \frac{\partial\bar{\Phi}_\nu}{\partial n}\right|^2 d\Gamma(\bar{f}_\nu)\right)^{\frac{1}{2}} \\
&\quad + \left(\int_{\Gamma(\bar{f}_\nu)}\left|\frac{\partial\bar{\Phi}_\nu}{\partial n}\right|^2 d\Gamma(\bar{f}_\nu)\right)^{\frac{1}{2}}\left(\int_{\Gamma(\bar{f}_\nu)} |\phi_{\bar{f}_\nu,\Psi} - \phi_{\bar{f}_\nu,\bar{\Phi}_\nu}|^2\,d\Gamma(\bar{f}_\nu)\right)^{\frac{1}{2}} \\
&\quad + \mathrm{Re}\int_{\Gamma(\bar{f}_\nu)} \phi_{\bar{f}_\nu,\bar{\Phi}_\nu}\frac{\partial\bar{\Phi}_\nu}{\partial n}\,d\Gamma(\bar{f}_\nu) + \nu\left(\left\|\frac{\partial\Psi}{\partial n} - \frac{\partial\bar{\Phi}_\nu}{\partial n}\right\|^2_{L^2(\Gamma(\bar{f}_\nu))} + \left\|\frac{\partial\bar{\Phi}_\nu}{\partial n} - G\right\|^2_{L^2(\Gamma(\bar{f}_\nu))}\right) \\
&\leq C\left\|\frac{\partial\Psi}{\partial n} - \frac{\partial\bar{\Phi}_\nu}{\partial n}\right\|^2_{L^2(\Gamma(\bar{f}_\nu))} + \left(M\int_{\Gamma(\bar{f}_\nu)} d\Gamma(\bar{f}_\nu)\right)\|\phi_{\bar{f}_\nu,\Psi} - \phi_{\bar{f}_\nu,\bar{\Phi}_\nu}\|_{L^2(\Gamma(\bar{f}_\nu))} \\
&\quad + \mathrm{Re}\int_{\Gamma(\bar{f}_\nu)} \phi_{\bar{f}_\nu,\bar{\Phi}_\nu}\frac{\partial\bar{\Phi}_\nu}{\partial n}\,d\Gamma(\bar{f}_\nu) + \nu\left(\left\|\frac{\partial\Psi}{\partial n} - \frac{\partial\bar{\Phi}_\nu}{\partial n}\right\|^2_{L^2(\Gamma(\bar{f}_\nu))} + \left\|\frac{\partial\bar{\Phi}_\nu}{\partial n} - G\right\|^2_{L^2(\Gamma(\bar{f}_\nu))}\right)
\end{aligned}$$

for suitable constants $C$ and $M$.

Choosing $\Psi = \bar{\Psi}^{(N_k)} \in S^{N_k}$ so that $\partial\bar{\Psi}^{(N_k)}/\partial n$ best approximates $\partial\bar{\Phi}_\nu/\partial n$, it follows that

$$\lim_{k\to\infty} \left\| \frac{\partial \bar{\Psi}^{(N_k)}}{\partial n} - \frac{\partial \bar{\Phi}_\nu}{\partial n} \right\|^2_{L^2(\Gamma(\bar{f}_\nu))} = 0,$$

while, from Lemma 4.1, we also have

$$\lim_{k\to\infty} \int_{\Gamma(\bar{f}_\nu)} |\phi_{\bar{f}_\nu \bar{\Psi}^{N_k}} - \phi^2_{\bar{f}_\nu \bar{\Phi}_\nu}| \, d\Gamma(\bar{f}_\nu) = 0.$$

It follows that

$$L_\nu^{(N_k)}(f_\nu^{(N_k)}, \Phi_\nu^{(N_k)}) \leq \operatorname{Re} \int_{\Gamma(\bar{f}_\nu)} \phi_{\bar{f}_\nu \bar{\Phi}_\nu} \frac{\partial \bar{\Phi}_\nu}{\partial n} d\Gamma(\bar{f}_\nu) + \nu \left\| \frac{\partial \bar{\Phi}_\nu}{\partial n} - g \right\|^2_{L^2(\Gamma(\bar{f}_\nu))} + o(1)$$

$$= L_\nu(\bar{f}_\nu, \bar{\Phi}_\nu) + o(1),$$

where $o(1) \to 0$ as $k \to \infty$. Since $L_\nu^{(N_k)}(f_\nu^{(N_k)}, \Phi_\nu^{(N_k)}) = L_\nu(f_\nu^{(N_k)}, \Phi_\nu^{(N_k)})$, continuity implies that

$$L_\nu(f_\nu^0, \Phi_\nu^0) \leq L_\nu(\bar{f}_\nu, \bar{\Phi}_\nu),$$

which completes the proof.

Finally, we address the question of the construction of a complete family of solutions. Recall that the class of admissible surfaces is defined in such a way that all contain a ball of radius $\frac{1}{2}a$, centred at the point $p_0$, for some preassigned constant $a > 0$. Certainly we may consider a surface $\Gamma_{a_0}$ strictly interior to the surface of the ball $\Gamma_{\frac{1}{2}a}$. We may then prove the following result.

THEOREM 4.2 *Let $\{\phi_n\}_{n=1}^\infty$ be a linearly independent family of functions that is complete in $L^2(\Gamma_{a_0})$. For each $n$, define the function $u_n$ by*

$$u_n(p) = \int_{\Gamma_{a_0}} \gamma(p, q)\phi_n(q) \, d\Gamma_q, \quad \text{for } p \in \mathbb{R}^3 \backslash D^-_{a_0}.$$

*Then the functions*

$$v_n(p) = (\partial u_n/\partial n)(p) \quad \text{for } p \in \Gamma(f), \quad n = 1, 2, \cdots \tag{4.12}$$

*form a complete linearly independent family in $L^2(\Gamma(f))$ for all $f \in U_{ad}$.*

*Remark.* By linear independence of these countable families we mean that any finite subset is linearly independent.

*Proof.* Consider the family $\{v_n\} \subset L^2(\Gamma(f))$ and suppose that there is some $\Psi \in L^2(\Gamma(f))$ such that

$$0 = \langle \Psi, v_n \rangle_{L^2(\Gamma(f))} \quad \text{for all } n = 1, 2, \cdots.$$

Then, recognizing that $\Gamma(f) \cap \Gamma_{a_0} = \emptyset$,

$$\begin{aligned} 0 = \langle \Psi, v_n \rangle_{L^2(\Gamma(f))} &:= \int_{\Gamma(f)} \Psi(p) \Big( \frac{\partial}{\partial n_p} \int_{\Gamma_{a_0}} \bar{\gamma}(p, q) \overline{\phi_n}(q) \, d\Gamma_q \Big) \, d\Gamma_p \\ &= \int_{\Gamma_{a_0}} \Big( \int_{\Gamma(f)} \Psi(p) \frac{\partial}{\partial n_p} \bar{\gamma}(p, q) \, d\Gamma_p \Big) \bar{\phi}_n(q) \, d\Gamma_q \\ &= \int_{\Gamma_{a_0}} (\bar{W}_{\Gamma(f)} \Psi)(q) \bar{\phi}_n(q) \, d\Gamma_q \\ &= \langle \bar{W}_{\Gamma(f)} \Psi, \phi_n \rangle_{L^2(\Gamma_{a_0})}, \end{aligned}$$

where $W_{\Gamma(f)}$ is the double layer operator defined implicitly above with the property that $\bar{W}_{\Gamma(f)} : L^2(\Gamma(f)) \to L^2(\Gamma_{a_0})$ and that the set $\{\phi_n\}$ is the complete set $L^2(\Gamma_{a_0})$. Therefore $\bar{W}_{\Gamma(f)} \Psi = 0$, and hence $W_{\Gamma(f)} \bar{\Psi} = 0$, on $\Gamma_{a_0}$.

Now consider the function $u_\Psi$, defined in $D_f^-$ by

$$u_\Psi(q) := \int_{\Gamma(f)} \frac{\partial}{\partial n_p} \gamma(p, q) \bar{\Psi}(p) \, d\Gamma_p.$$

Then $u_\Psi$ is harmonic in $D_f^-$. Moreover, in the region $D_{a_0}^-$, i.e. in the region interior to $\Gamma_{a_0}$, $u_\Psi$ is harmonic and $u_\Psi|_{\Gamma_{a_0}} = 0$. Thus $u_\Psi$ vanishes everywhere in $D_{a_0}^-$ and hence, by analytic continuation, everywhere in $D_f^-$. But, since $u_\Psi = W_{\Gamma(f)} \bar{\Psi}$, we have, using the jump conditions,

$$0 = \lim_{q \to \Gamma(f)^-} W_{\Gamma(f)} \bar{\Psi}(q) = (I + \bar{K}^*_{\Gamma(f)}) \bar{\Psi}.$$

The results of Section 2 guarantee that the only solution of

$$(I + \bar{K}^*_{\Gamma(f)}) u = 0$$

is the trivial solution, so we conclude that $\Psi = 0$. This establishes completeness.

To establish that any finite subset of the $v_n$ is linearly independent, suppose, without loss of generality, that there exist constants $\alpha_1, \ldots, \alpha_N$ such that

$$\sum_{i=1}^{N} \alpha_i v_i = 0 \quad \text{on } \Gamma(f).$$

Then, by definition,

$$\begin{aligned} 0 = \sum_{i=1}^{N} \alpha_i v_i &= \frac{\partial}{\partial n} \sum_{i=1}^{N} \alpha_i \int_{\Gamma_{a_0}} \gamma(p, q) \phi_i(q) \, d\Gamma_q \\ &= \frac{\partial}{\partial n} \int_{\Gamma_{a_0}} \gamma(p, q) \sum_{i=1}^{N} \alpha_i \phi_i(q) \, d\Gamma_q \\ &= \frac{\partial}{\partial n} \int_{\Gamma_{a_0}} \gamma(p, q) w(q) \, d\Gamma_q \quad \text{for } p \in \Gamma(f), \end{aligned}$$

where $w(q) := \sum_{i=1}^{N} \alpha_i \phi_i(q)$.

Now, for $p \in D_{a_0}^+$, i.e. in the exterior $\Gamma_{a_0}$, define the function $v = v(p)$ by

$$v(p) := \int_{\Gamma_{a_0}} \gamma(p, q) w(q) \, d\Gamma_q = S_{a_0} w.$$

Thus certainly $v$ is harmonic in $D_{a_0}^-$ and, in particular, is harmonic in $D_f^-$. Moreover $v$ satisfies the radiation condition (3.3e) and the boundary condition

$$\frac{\partial}{\partial n} v\Big|_{\Gamma(f)} = 0,$$

and so $v$ is a solution of the exterior homogeneous Neumann problem. By the uniqueness theorem, $v$ vanishes identically in $D_f^+$ and so, by analytic continuation, also vanishes in the region exterior to $\Gamma_{a_0}$. Hence, again using the jump conditions,

$$0 = \lim_{p \to \Gamma_{a_0}} \frac{\partial}{\partial n} S_{a_0} w = (I + K_{a_0}) w = 0,$$

where

$$(K_{a_0} w)(p) := \int_{\Gamma_{a_0}} \frac{\partial \gamma}{\partial n_p}(p, q) w(q)\, \mathrm{d}\Gamma_q.$$

But $(I + \bar{K}_{a_0}^*)u = 0$ is uniquely solvable, so $(I + \bar{K}_{a_0})v = 0$ is also uniquely solvable by the Fredholm alternative. Thus, given $(I + K_{a_0})w = 0$, it follows that $(I + \bar{K}_{a_0})\bar{w} = 0$, and so $\bar{w}$ vanishes on $\Gamma_{a_0}$. Therefore, so does $w$. From the linear independence of the $\{\phi_i\}_{i=1}^N$, we conclude that all the coefficients $\alpha_i = 0$ $(i = 1, \ldots, N)$, and this shows that the corresponding functions $v_i$ $(i = 1, \ldots, N)$ are likewise linearly independent.

**Acknowledgement**

This work was supported under NSF Grant No. DMS-8912593 and AFOSR Grant 86-0269.

## References

Angell, T. S., & Kleinman, R. E. 1984 A Galerkin procedure for optimization in radiation problems. *SIAM J. Appl. Math.* **44**, 1246–57.

Angell, T. S., & Kleinman, R. E. 1985 A new optimization method for antenna design. *Ann. Télécommun.* **40**, 341–9.

Angell, T. S., Hsiao, G. C., & Kleinman, R. E. 1986 An optimal design problem for submerged bodies. *Math. Meth. Appl. Sci.* **8**, 50–76.

Angell, T. S., Kleinman, R. E., & Roach, G. F. 1987 An inverse transmission problem for the Helmholtz equation. *Inverse Problems* **3**, 149–80.

Cea, J. 1975 Identification de domaines. *Lecture Notes in Computer Science* 3. Berlin: Springer, pp. 92–102.

Cea, J., Gioan, A., & Michel, J. 1974 Adaptation de la méthode du gradient à un probléme d'identification de domaine. *Lecture Notes in Computer Science* 11. Berlin: Springer.

Chenais, D. 1975 On the existence of a solution in a domain identification problem. *J. Math. Anal. Appl.* **52**, 189–209.

Colton, D., & Kress, R. 1983 *Integral Equation Methods in Scattering Theory*. New York: Wiley.

Hulme, A. 1984 Some applications of Maz'ja's uniqueness theorem to a class of linear water wave problems. *Math. Proc. Camb. Phil. Soc.* **95**, 165–74.

JOHN, F. 1950 On the motion of floating bodies II. *Commun. Pure Appl. Math.* **3**, 45–101.

KIRSCH, A., & KRESS, R. 1986 On an integral equation of the first kind in inverse acoustic scattering. In: *Inverse Problems* (J. R. Cannon & U. Horning, eds.), ISNM 77. Basel: Birkhaüser, pp. 93–102.

KLEINMAN, R. E. 1982 On the mathematical theory of the motion of floating bodies—an update. David W. Taylor Naval Ship Research and Development Center Report DTNSRDC-82/074, Bethesda, MD.

LUENBERGER, D. G. 1969 *Optimization by Vector Space Methods.* New York: Wiley.

LIONS, J. L. 1971 *Optimal Control of Systems Governed by Partial Differential Equations* (S. K. Mitter, tr.). Berlin: Springer.

LIONS, J. L. 1972 *Some Aspects of the Optimal Control of Distributed Parameter Systems.* Regional Conference Series in Applied Mathematics, SIAM.

MAZ'JA, V. G. 1978 Solvability of the problem on the oscillations of a fluid containing a submerged body. *J. Soviet Math.* **10**, 86–9.

MIRANDA, C. 1970 *Partial Differential Equations of Elliptic Type,* 2nd rev. edn. Berlin: Springer.

NEWMAN, J. N. 1977 *Marine Hydrodynamics.* Cambridge, MA: MIT Press.

PIRONNEAU, O. 1973 On optimum profiles in Stokes flow. *J. Fluid Mech.* **59**, 117–28.

PIRONNEAU, O. 1974 On optimal design in fluid mechanics. *J. Fluid Mech.* **64**, 97–110.

SIMON, N. J., & URSELL, F. 1984 Uniqueness in linearized two-dimensional water-wave problems. *J. Fluid Mech.* **148**, 137–54.

WEHAUSEN, J. V., & LAITONE, E. V. 1960 Surface waves. In: *Handbuch der Physik,* Bd. IX: *Strömungsmechanik* III (S. Flügge, ed.). Berlin: Springer. pp. 446–778.

SIAM J. APPL. MATH.
Vol. 52, No. 6, pp. 1597–1610, December 1992

006

# THE CONDUCTIVE BOUNDARY CONDITION FOR MAXWELL'S EQUATIONS*

T. S. ANGELL† AND A. KIRSCH‡

**Abstract.** First, the conductive boundary value problem is derived for the quasi-stationary Maxwell equations that arise in the study of magnetotellurics. Then the boundary integral equation method is used to prove the existence and uniqueness of solutions of the problem. The final section is devoted to a study of the set of far field patterns for scattering problems with plane wave incidence.

**Key words.** Maxwell equations, boundary integral equations, scattering theory

**AMS(MOS) subject classifications.** 35P25, 45B05, 78A45, 86A20

**1. Introduction.** Geophysicists, in their study of electromagnetic induction in the earth (called magnetotellurics) commonly use a boundary condition for the electromagnetic field, which is often referred to as the *conductive boundary condition*. We refer to Schmucker [16] or Vasseur and Weidelt [19] for the physical explanation of this boundary condition. This boundary condition models the occurrence of a thin layer of very high conductivity for, while it is well known that the electric field does not penetrate into an ideal conductor of positive thickness, such a field certainly will penetrate into the medium beyond that conductor if the latter is infinitely thin.

The analogous boundary condition in both the electromagnetic and acoustic problems have been known for some time; see, e.g., Harrington and Mautz [6] or Senior [17]. In this context, the conditions have been considered as approximations to the full transmission conditions. The wellposedness of the boundary value problem in the scalar case has only recently been treated by Hettlich [7] and Angell, Kleinman, and Hettlich [1]. In this paper, we employ the technique of boundary integral equations to discuss the existence of solutions to the electromagnetic conductive problem.

The use of integral equations in problems of acoustics and electromagnetics is a well-known technique; a current account of the method may be found in Colton and Kress [4]. For the problem of scattering of time harmonic electromagnetic waves by a perfectly conducting object, the method was applied at least as early as 1949 by Maue [13]. Müller [14] used the method in 1951 to treat the electromagnetic transmission problem. As the classical transmission conditions are a special case of the boundary conditions of the problem discussed here, and our surfaces may be less regular than those of [14], the present work may be considered as a generalization of Müller's results.

Without attempting to give an exhaustive review of the literature on integral equations in electromagnetics, we mention that various aspects have also been treated by Weyl [22], Saunders [15], Calderòn [2], Werner [20], [21], Knauff and Kress [9], and Gray and Kleinman [5]. More recently, Marx [11], [12] has developed a single equation for electromagnetic and time-dependent scattering problems.

In §2 we derive the conductive boundary condition from the quasi-stationary Maxwell equations for induction problems in a layered half-space. This is the situation

* Received by the editors March 26, 1990; accepted for publication (in revised form) November 12, 1990.

† Department of Mathematical Sciences, University of Delaware, Newark, Delaware 19716. This author was supported by Air Force Office of Scientific Research grant AFOSR-91-0277.

‡ Institut für Angewandte Mathematik, Universität Erlangen–Nürnberg, D-8520 Erlangen, Germany. This author was supported by Deutsche Forschungsgemeinschaft (DFG) grant KR940/1-1.

of magnetotellurics. In §3 we consider the model problem where the anomalous region of conductivity is embedded in a homogeneous "full space." We will prove uniqueness and existence results for classical solutions.

We devote §4 to the description of the far field patterns for scattering problems whose incident fields are given by plane wave solutions of the Maxwell equations. We prove the reciprocity principle and use this result to show that the class of all far field patterns corresponding to the incident plane waves of any direction and amplitude is dense in $L_T^2(S^2)$, the space of all $L^2$-tangential fields on the unit sphere $S^2$, provided that the pair $(k_1, k_2)$ of wavenumbers is not an eigenvalue of a related eigenvalue problem.

**2. Physical derivation of a conductive boundary condition in magnetotellurics.** We model the earth as a layered half-space filling the region $x_3 \geq 0$. The conductivity $\sigma_n$ (normal conductivity) of the earth is assumed to depend only on depth $x_3$ and to be piecewise constant.

A bounded region $\Omega$ (anomalous region) is imbedded in the half-space $x_3 > 0$. The conductivity $\sigma_a$ in $\Omega$ is different from $\sigma_n$ and is allowed to depend on $\mathbf{x} = (x_1, x_2, x_3)$. Furthermore, we assume that $\Omega$ is covered by a thin layer with very high conductivity $\sigma_l^\epsilon = \sigma_l^\epsilon(\mathbf{x})$, such that the integrated conductivity

$$\tau(\mathbf{x}) := \lim_{\epsilon \to 0+} \int_{-\epsilon}^{\epsilon} \sigma_l^\epsilon(\mathbf{x} + t\mathbf{n}(\mathbf{x}))\, dt, \qquad \mathbf{x} \in \partial\Omega,$$

remains finite, i.e., $\sigma_l(\mathbf{x} + t\mathbf{n}(\mathbf{x})) = \tau(\mathbf{x})\delta(t)$. Here we denote by $\mathbf{n}(\mathbf{x})$ the outer unit normal vector at $\mathbf{x} \in \partial\Omega$.

We now assume that some kind of sources in the half-space $x_3 < 0$ (e.g., in the ionosphere) induces an electromagnetic field $\mathbf{E}$, $\mathbf{B}$ in the earth $x_3 > 0$. Here $\mathbf{E} = \mathbf{E}(\mathbf{x})$, and $\mathbf{B} = \mathbf{B}(\mathbf{x})$ denote the spacial parts of the electric field $\mathbf{E}(\mathbf{x})e^{-i\omega t}$ and magnetic field $\mathbf{B}(\mathbf{x})e^{-i\omega t}$, where $\omega > 0$ denotes the frequency. Then $\mathbf{E}$ and $\mathbf{B}$ satisfy Maxwell equations in $x_3 > 0$. We formulate them in their quasi-stationary approximation although this is not necessary for the mathematical theory, as follows:

$$\operatorname{curl} \mathbf{B} = \mu_0 \sigma \mathbf{E}, \qquad \operatorname{curl} \mathbf{E} = i\omega \mathbf{B}, \tag{2.1}$$

where $\sigma$ is the conductivity ($\sigma = \sigma_n$ for $\mathbf{x} \notin \bar{\Omega}$, $\sigma = \sigma_a$ for $\mathbf{x} \in \Omega$) and $\mu_0$ the magnetic permeability in vacuum. Using SI units throughout, we measure $\mathbf{E}$ in V/m, $\mathbf{B}$ in Tesla = $\mathrm{Vs/m^2}$, $\sigma$ in A/Vm, $\omega$ in 1/s, and $\mu_0 = 4\pi 10^{-7}$Vs/Am.

To derive the boundary conditions, let the layer with conductivity $\sigma_l^\epsilon$ be of finite thickness $\epsilon > 0$. Let $C$ be a $C^2$-arc on $\partial\Omega$ with unit tangential vector $\ell(\mathbf{x})$, $\mathbf{x} \in C$, and

$$S = \{\mathbf{x} + t\mathbf{n}(\mathbf{x}) : \mathbf{x} \in C, |t| < \epsilon\}$$

the surface perpendicular to $\partial\Omega$ with boundary $\partial S$.

For $\mathbf{x} \in C$, let $\nu(\mathbf{x}) = \mathbf{n}(\mathbf{x}) \times \ell(\mathbf{x})$, where $\mathbf{a} \times \mathbf{b}$ and $\mathbf{a} \cdot \mathbf{b}$ are mean vector products and scalar products, respectively. Then the Stokes theorem yields

$$\int_{\partial S} \mathbf{E} \cdot d\ell = \int_S \nu \cdot \operatorname{curl} \mathbf{E}\, ds = i\omega \int_S \nu \cdot \mathbf{B}\, ds \tag{2.2}$$

and

$$(2.3)\quad \int_{\partial S} \mathbf{B}\cdot d\ell = \int_S \nu\cdot \operatorname{curl}\mathbf{B}\, ds = \mu_0 \int_S \sigma_l^\epsilon\, \nu\cdot\mathbf{E}\, ds$$
$$= \mu_0 \int_C \int_{-\epsilon}^{+\epsilon} \sigma_l^\epsilon(\mathbf{x}+t\mathbf{n}(\mathbf{x}))\nu(\mathbf{x})\cdot\mathbf{E}(\mathbf{x}+t\mathbf{n}(\mathbf{x}))(1+\mathcal{O}(\epsilon))\, dt\, d\ell(\mathbf{x}).$$

For $\epsilon\to 0$, we conclude from (2.2) that $\int_C(\mathbf{E}_+ - \mathbf{E}_-)\cdot d\ell = 0$. Here $\mathbf{E}_\pm$ denotes the limit of $\mathbf{E}$ from the outside $(+)$ and the inside $(-)$, respectively. By the mean value theorem for integrals and with $\epsilon\to 0$ in (2.3), we arrive at

$$\int_C(\mathbf{B}_+ - \mathbf{B}_-)\cdot d\ell = \mu_0\int_C \tau\nu\cdot\mathbf{E}\, d\ell.$$

This holds for every arc $C$; thus

$$(2.4)\quad \begin{aligned} \mathbf{n}\times\mathbf{E}|_+ - \mathbf{n}\times\mathbf{E}|_- &= 0 \quad \text{on } \partial\Omega,\\ \mathbf{n}\times\mathbf{B}|_+ - \mathbf{n}\times\mathbf{B}|_- &= \mu_0\tau(\mathbf{n}\times\mathbf{E})\times\mathbf{n} \quad \text{on } \partial\Omega. \end{aligned}$$

It is the aim of this paper to study (2.1), (2.4), together with a radiation condition for the "anomalous" parts of $\mathbf{E}$ and $\mathbf{B}$ for the special case of a homogeneous region in a homogeneous full-space (i.e., $\sigma_n$ and $\sigma_a$ are constant).

Thus let us assume that $\sigma_n$ and $\sigma_a$ are constant. First, we symmetrize (2.1) and define $k_1^2 = i\omega\mu_0\sigma_a$, $k_2^2 = i\omega\mu_0\sigma_n$ such that $\operatorname{Im} k_j > 0$, $j = 1,2$, and $\mathbf{H} = (\omega/k_1)\mathbf{B}$ in $\Omega$, $\mathbf{H} = (\omega/k_2)\mathbf{B}$ in $\mathbf{R}^3\setminus\overline{\Omega}$. Then we see that (2.1) takes the form

$$(2.5)\qquad \operatorname{curl}\mathbf{H} = -ik\mathbf{E}, \qquad \operatorname{curl}\mathbf{E} = ik\mathbf{H} \quad \text{in } \mathbf{R}^3\setminus\partial\Omega,$$

with

$$k(\mathbf{x}) = \begin{cases} k_1, & \text{in } \Omega,\\ k_2, & \text{in } \mathbf{R}^3\setminus\overline{\Omega}. \end{cases}$$

The boundary conditions (2.4) change into

$$(2.6)\quad \begin{aligned} \mathbf{n}\times\mathbf{E}|_+ - \mathbf{n}\times\mathbf{E}|_- &= 0 \quad \text{on } \partial\Omega,\\ \mathbf{n}\times[\mathbf{n}\times(k_2\mathbf{H}|_+ - k_1\mathbf{H}|_-)] &= \mu_0\tau\omega\,\mathbf{n}\times\mathbf{E} \quad \text{on } \partial\Omega. \end{aligned}$$

**3. Uniqueness and existence of solutions for a model problem.** We now focus on the main problem of this paper. Given an open and bounded region $\Omega\subset\mathbf{R}^3$ with $C^2$-boundary $\partial\Omega$, numbers $k_1, k_2, \mu_1, \mu_2 \in \mathbf{C}\setminus\{0\}$ with $\operatorname{Im} k_j \geq 0\,(j=1,2)$, a complex-valued function $\beta\in C^{0,\alpha}(\partial\Omega)$, a direction $\hat{\mathbf{d}}\in S^2$, and an amplitude $\mathbf{p}\in\mathbf{C}^3$ with $\hat{\mathbf{d}}\cdot\mathbf{p} = 0$, find vector fields $\mathbf{E},\mathbf{H}\in C^{1,\alpha}(\mathbf{R}^3\setminus\partial\Omega)\cap C'$, which satisfy

$$(3.1)\qquad \operatorname{curl}\mathbf{E} - ik\mathbf{H} = 0,\ \operatorname{curl}\mathbf{H} + ik\mathbf{E} = 0 \quad \text{in } \mathbf{R}^3\setminus\partial\Omega;$$

$$(3.2)\quad \begin{aligned} \mathbf{n}\times\mathbf{E}|_+ - \mathbf{n}\times\mathbf{E}|_- &= 0 \quad \text{on } \partial\Omega;\\ \mu_2\mathbf{n}\times(\mathbf{n}\times\mathbf{H})|_+ - \mu_1\mathbf{n}\times(\mathbf{n}\times\mathbf{H})|_- &= \beta\mathbf{n}\times\mathbf{E} \quad \text{on } \partial\Omega; \end{aligned}$$

$$(3.3)\qquad \mathbf{E}(\mathbf{x}) = \mathbf{E}^i(\mathbf{x}) + \mathbf{E}^s(\mathbf{x}),\ \mathbf{H}(\mathbf{x}) = \frac{1}{ik_2}\operatorname{curl}\mathbf{E}^i(\mathbf{x}) + \mathbf{H}^s(\mathbf{x}), \qquad \mathbf{x}\notin\overline{\Omega},$$

with incident field $\mathbf{E}^i(\mathbf{x}) = \mathbf{p}e^{ik_2\hat{\mathbf{d}}\cdot\mathbf{x}}$, and where the scattered fields $\mathbf{E}^s$ and $\mathbf{H}^s$ satisfy the radiation condition

$$(3.4)\quad \frac{\mathbf{x}}{|\mathbf{x}|}\times\mathbf{H}^s(\mathbf{x}) + \mathbf{E}^s(\mathbf{x}) = o\left(\frac{1}{|\mathbf{x}|}\right), \quad |\mathbf{x}|\to\infty, \qquad \text{uniformly in } \frac{\mathbf{x}}{|\mathbf{x}|}\in S^2$$

with $k = k_1$ in $\Omega$ and $k = k_2$ in $\mathbf{R}^3 \setminus \overline{\Omega}$. Likewise, we will set $\mu = \mu_1$ in $\Omega$ and $\mu = \mu_2$ in $\mathbf{R}^3 \setminus \Omega$.

Here we have used the following notational conventions:

(i) $C^{0,\alpha}(\partial\Omega)$ denotes the space of Hölder continuous functions on $\partial\Omega$ of order $\alpha \in (0,1)$ (the space $C^{1,\alpha}(\partial\Omega)$ is defined analogously);

(ii) $C' = \{\mathbf{E} : \mathbf{R}^3 \setminus \partial\Omega \to \mathbf{C}^3 \,:\, \mathbf{E}|_\Omega \in C(\overline{\Omega}),\ \mathbf{E}|_{\mathbf{R}^3\setminus\overline{\Omega}} \in C(\mathbf{R}^3 \setminus \Omega)\}$;

(iii) $F|_{+(-)}$ denotes the limit of $F$ on $\partial\Omega$ from the exterior (interior);

(iv) $\mathbf{n}(\mathbf{x})$ denotes the outer unit normal vector at $\mathbf{x} \in \partial\Omega$, and $\mathbf{a} \times \mathbf{b}$, $\mathbf{a} \cdot \mathbf{b}$ are the vector products and scalar products, respectively.

The situation discussed in §2 is covered by setting

$$\begin{aligned} &k_1^2 = i\omega\mu_0\sigma_a, \quad k_2^2 = i\omega\mu_0\sigma_n, \qquad (\operatorname{Im} k_1 > 0,\ \operatorname{Im} k_2 > 0), \\ &\mu_j = k_j \ (j = 1,2), \quad \beta = \omega\mu_0\tau, \quad \mathbf{H} = \frac{\omega}{k}\mathbf{B}. \end{aligned} \tag{3.5}$$

THEOREM 3.1 (uniqueness). *Let the parameters of the problem satisfy the following relations:*

$$\operatorname{Re}(\beta\overline{\mu}_2) \geq 0 \text{ on } \partial\Omega, \quad \operatorname{Im}\left(k_1\frac{\mu_1}{\mu_2}\right) \geq 0, \quad \operatorname{Im}\left(\overline{k_1}\frac{\mu_1}{\mu_2}\right) \leq 0. \tag{3.6}$$

*Then there exists at most one solution* $(\mathbf{E}, \mathbf{H})$ *of problems* (3.1)–(3.4).

*Proof.* Let $\mathbf{E}^i = 0$, i.e., $(\mathbf{E}, \mathbf{H})$ satisfies (3.1), (3.2), and the radiation condition (3.4). We use Green's theorem for vector fields in $\Omega$ as follows:

$$\int_\Omega (\mathbf{E} \cdot \Delta\overline{\mathbf{E}} + \operatorname{curl}\mathbf{E} \cdot \operatorname{curl}\overline{\mathbf{E}})\, d\mathbf{x} = \int_{\partial\Omega} \mathbf{n} \cdot (\mathbf{E} \times \operatorname{curl}\overline{\mathbf{E}})\, ds.$$

Then, with $\operatorname{curl}\mathbf{E}_j = ik_j\mathbf{H}$ and $\Delta\mathbf{E}_j = -k_j^2 E$, $j = 1,2$ in $\Omega$ and $\Omega_R$, respectively, $\Omega_R := \{x \in \mathbf{R}^3\setminus\overline{\Omega} : |\mathbf{x}| < \mathbf{R}\}$, we may add and use the boundary conditions to obtain

$$\begin{aligned} &\int_{|\mathbf{x}|<R} \overline{(\mu/k)}(-\overline{k}^2|\mathbf{E}|^2 + |k\mathbf{H}|^2)\, d\mathbf{x} \\ &= -i\int_{\partial\Omega} \beta|\mathbf{n} \times \mathbf{E}|^2\, ds - i\overline{\mu}_2 \int_{|\mathbf{x}|=R} \mathbf{n} \cdot (\mathbf{E} \times \overline{\mathbf{H}})\, ds. \end{aligned} \tag{3.7}$$

From the radiation condition integrated over the sphere of radius $R$, we see that

$$\begin{aligned} &\int_{|\mathbf{x}|=R} (|\mathbf{n} \times \mathbf{H}|^2 + |\mathbf{E}|^2)\, ds - 2\operatorname{Re}\int_{|\mathbf{x}|=R} \mathbf{n} \cdot (\mathbf{E} \times \mathbf{H})\, ds \\ &= \int_{|\mathbf{x}|=R} |\mathbf{n} \times \mathbf{H} + \mathbf{E}|^2\, ds = o(1) \quad \text{for } R \to \infty, \end{aligned}$$

and, by dividing (3.7) by $\overline{\mu}_2$ and taking the imaginary part,

$$\begin{aligned} &\int_{|\mathbf{x}|<R} \operatorname{Im}(\mu k/\mu_2)\, |\mathbf{E}|^2\, d\mathbf{x} - \int_{|\mathbf{x}|<R} \operatorname{Im}(\mu/(\mu_2 k))\, |k\mathbf{H}|^2\, d\mathbf{x} \\ &= -\int_{\partial\Omega} \operatorname{Re}(\beta/\mu_2)\, |\mathbf{n} \times \mathbf{E}|^2\, ds - \frac{1}{2}\int_{|\mathbf{x}|=R} (|\mathbf{n} \times \mathbf{H}|^2 + |\mathbf{E}|^2)\, ds + o(1) \\ &\qquad\qquad \text{for } R \to \infty. \end{aligned}$$

From (3.6) we see that $\operatorname{Im}(\mu k/\mu_2) \geq 0$, $\operatorname{Im}(\mu/\mu_2 k) \leq 0$, and $\operatorname{Re}(\lambda/\mu_2) \geq 0$ on $\partial\Omega$. This implies that

$$\text{(3.8)} \qquad \operatorname{Im}(\mu k/\mu_2)\,\mathbf{E} \equiv 0 \text{ in } \mathbf{R}^3 \setminus \partial\Omega \quad \text{and} \quad \int_{|\mathbf{x}|=R} |\mathbf{E}|^2\, ds \to 0 \ (R \to \infty).$$

If $\operatorname{Im} k_2 = 0$, then, from Rellich's theorem (cf. [4]), it follows that $\mathbf{H} \equiv \mathbf{E} \equiv 0$ in $\mathbf{R}^3 \setminus \Omega$. If, on the other hand, $\operatorname{Im} k_2 > 0$, then the identities $\mathbf{H} \equiv \mathbf{E} \equiv 0$ in $\mathbf{R}^3 \setminus \Omega$ follow from the first identity in (3.8). In either case, $\mathbf{n} \times \mathbf{E}|_- = 0$ and $\mathbf{n} \times (\mathbf{n} \times \mathbf{H})|_- = 0$ on $\partial\Omega$, which implies, by the representation theorem, that $\mathbf{E} \equiv \mathbf{H} \equiv 0$ throughout $\Omega$, and the proof is complete.

We remark that, for the physically relevant situation described by relations (3.5), the uniqueness assumptions (3.6) are satisfied, as is readily verified.

Now we assume a layer ansatz and use the integral equation method (see [10] for some related boundary value problems) to prove existence of a solution to our model problem (3.1)–(3.4). First, we define the scalar three-dimensional fundamental solution corresponding to $k_j$ $(j = 1, 2)$ as

$$\Phi_j(\mathbf{x}, \mathbf{y}) := \frac{e^{ik_j|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}, \qquad \mathbf{x} \neq \mathbf{y}, \quad j = 1, 2$$

and set

$$\Phi(\mathbf{x}, \mathbf{y}) = \begin{cases} \Phi_1(\mathbf{x}, \mathbf{y}) & \text{if } \mathbf{x} \in \Omega, \mathbf{y} \in \partial\Omega, \\ \Phi_2(\mathbf{x}, \mathbf{y}) & \text{if } \mathbf{x} \notin \overline{\Omega}, \mathbf{y} \in \partial\Omega; \end{cases}$$

$$k(\mathbf{x}) = \begin{cases} k_1 & \text{if } \mathbf{x} \in \Omega, \\ k_2 & \text{if } \mathbf{x} \notin \overline{\Omega}; \end{cases} \qquad \mu(\mathbf{x}) = \begin{cases} \mu_1 & \text{if } \mathbf{x} \in \Omega, \\ \mu_2 & \text{if } \mathbf{x} \notin \overline{\Omega}. \end{cases}$$

We make an ansatz for $\mathbf{E}^s$, $\mathbf{H}^s$ in the form of a sum of electric and magnetic dipoles distributed on the boundary surface

$$\text{(3.9)} \qquad \begin{aligned} \mathbf{E}^s(\mathbf{x}) &= k(\mathbf{x})\operatorname{curl} \int_{\partial\Omega} \mathbf{a}(\mathbf{y})\Phi(\mathbf{x}, \mathbf{y})\, ds(\mathbf{y}) + \operatorname{curl}^2 \int_{\partial\Omega} \mathbf{b}(\mathbf{y})\Phi(\mathbf{x}, \mathbf{y})\, ds(\mathbf{y}), \\ \mathbf{H}^s(\mathbf{x}) &= \frac{1}{ik}\operatorname{curl} \mathbf{E}^s(\mathbf{x}) \end{aligned}$$

in $\mathbf{R}^3 \setminus \partial\Omega$, where $\mathbf{a}, \mathbf{b} \in C_D^{0,\alpha} := \{\mathbf{c} \in C_T^{0,\alpha} : \operatorname{Div} \mathbf{c} \in C^{0,\alpha}(\partial\Omega)\}$ with $C_T^{0,\alpha} := \{\mathbf{c} \in C^{0,\alpha}(\partial\Omega, \mathbf{C}^3) : \mathbf{n} \cdot \mathbf{c} = 0 \text{ on } \partial\Omega\}$ are unknown vector fields. By $\operatorname{Div} \mathbf{c}$, we denote the surface divergence of $\mathbf{c} \in C_T^{0,\alpha}$ (cf. [4, p. 60] for a definition).

From the properties of $\Phi_j$, we see that $(\mathbf{E}^s, \mathbf{H}^s)$ satisfies the Maxwell equations (3.1) in $\mathbf{R}^3 \setminus \partial\Omega$ and the radiation condition (3.4). By standard arguments (cf. [4, §2.6], we can show that $\mathbf{E}^s$ and $\mathbf{H}^s$ belong to $C'$.

The tangential components of $\mathbf{E}^s$ and $\mathbf{H}^s$ on $\partial\Omega$ take the form

$$\begin{aligned} \mathbf{n} \times \mathbf{E}^s|_\pm &= k_j \left[ \pm\frac{1}{2}\mathbf{a} + \mathbf{n} \times \int_{\partial\Omega} \operatorname{curl}_x(\mathbf{a}\Phi_j)\, ds(\mathbf{y}) \right] \\ &\quad + \mathbf{n} \times \operatorname{curl}^2 \int_{\partial\Omega} \mathbf{b}\Phi_j\, ds(\mathbf{y}), \qquad \mathbf{x} \in \partial\Omega, \\ \mathbf{n} \times \mathbf{H}^s|_\pm &= \frac{k_j}{i} \left[ \pm\frac{1}{2}\mathbf{b} + \mathbf{n} \times \int_{\partial\Omega} \operatorname{curl}_x(\mathbf{b}\Phi_j)\, ds(\mathbf{y}) \right] \\ &\quad + \frac{1}{i}\mathbf{n} \times \operatorname{curl}^2 \int_{\partial\Omega} \mathbf{a}\Phi_j\, ds(\mathbf{y}), \qquad \mathbf{x} \in \partial\Omega, \end{aligned}$$

the upper sign + and number $j = 2$ corresponding to the limit from $\mathbf{R}^3 \setminus \overline{\Omega}$, and the lower sign $-$ and $j = 1$ to that from the inside $\Omega$. These jump conditions follow from that for the curl of the single layer potential, which is proved in [4, Thm. 2.26].

The boundary conditions (3.2) for $\mathbf{E}^s$ and $\mathbf{H}^s$ lead to a system of integral equations on $\partial\Omega$ for the unknowns $\mathbf{a}$ and $\mathbf{b}$. Before we write them, we introduce the following boundary operators:

$$(Q\mathbf{b})(\mathbf{x}) = \mathbf{n}(\mathbf{x}) \times \mathbf{b}(\mathbf{x}), \qquad \mathbf{x} \in \partial\Omega,$$

$$(M_j\mathbf{b})(\mathbf{x}) = \mathbf{n}(\mathbf{x}) \times \int_{\partial\Omega} \operatorname{curl}_x(\mathbf{b}(\mathbf{y})\Phi_j(\mathbf{x},\mathbf{y}))\, ds(\mathbf{y}), \qquad \mathbf{x} \in \partial\Omega, \quad j = 1,2,$$

$$(P_j\mathbf{b})(\mathbf{x}) = \mathbf{n}(\mathbf{x}) \times \operatorname{curl}^2 \int_{\partial\Omega} \mathbf{b}(\mathbf{y})\Phi_j(\mathbf{x},\mathbf{y})\, ds(\mathbf{y}), \qquad \mathbf{x} \in \partial\Omega, \quad j = 1,2.$$

Then (3.2) leads to

$$\frac{1}{2}(k_1 + k_2)\mathbf{a} + (k_2M_2 - k_1M_1)\mathbf{a} + P_2\mathbf{b} - P_1\mathbf{b} = 0, \tag{3.10}$$

$$\begin{aligned}\left(\frac{\mu_2 k_2}{2i} + \frac{\mu_1 k_1}{2i}\right) Q\mathbf{b} + \frac{1}{i}(\mu_2 k_2 QM_2 - \mu_1 k_1 QM_1)\mathbf{b} \\ + \frac{1}{i}Q(\mu_2 P_2 - \mu_1 P_1)\mathbf{a} - \beta\left[-\frac{1}{2}k_1\mathbf{a} + k_1 M_1\mathbf{a} + P_1\mathbf{b}\right] = \mathbf{d},\end{aligned} \tag{3.11}$$

where

$$\begin{aligned}\mathbf{d}(\mathbf{x}) = (\mu_1 - \mu_2)\mathbf{n}(\mathbf{x}) \times \left[\mathbf{n}(\mathbf{x}) \times (\hat{\mathbf{d}} \times \mathbf{p})e^{ik_2\hat{\mathbf{d}}\cdot\mathbf{x}}\right] \\ + \beta(\mathbf{x})\mathbf{n}(\mathbf{x}) \times \mathbf{p}e^{ik_2\hat{\mathbf{d}}\cdot\mathbf{x}}, \qquad \mathbf{x} \in \partial\Omega.\end{aligned} \tag{3.12}$$

Thus we have the following theorem.

THEOREM 3.2. *The vector functions* $\mathbf{a}, \mathbf{b} \in C_D^{0,\alpha}$ *are solutions to system* (3.10), (3.11) *with* $\mathbf{d}$ *given by* (3.12) *if and only if the fields* $(\mathbf{E}^s, \mathbf{H}^s)$ *from* (3.9) *solve the boundary value problem* (3.1)–(3.4).

To discuss the solvability of (3.10), (3.11), we write them in matrix form as follows:

$$(L + K)\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix} = \begin{pmatrix}0\\ \mathbf{d}\end{pmatrix}, \tag{3.13}$$

where

$$L = \begin{pmatrix} \frac{1}{2}(k_1 + k_2)I, & 0 \\ \frac{1}{i}Q(\mu_2 P_2 - \mu_1 P_1) + \frac{\beta k_1}{2}I, & \frac{\alpha}{2i}Q - \beta P_1 \end{pmatrix},$$

$$K = \begin{pmatrix} k_2M_2 - k_1M_1, & P_2 - P_1 \\ -\lambda k_1 M_1, & \frac{1}{i}(\mu_2 k_2 QM_2 - \mu_1 k_1 QM_1) \end{pmatrix},$$

and $\alpha = k_1\mu_1 + k_2\mu_2$. We treat this equation in the spaces

$$L + K : C_D^{0,\alpha} \times C_D^{0,\alpha} \longrightarrow C_D^{0,\alpha} \times C_T^{0,\alpha}.$$

It is known that $K$ is compact between these spaces (cf. [4]) and that $L$ is well defined and bounded.

Our aim is to apply the Riesz theory for compact operators; we first show that a certain compact perturbation of $\frac{1}{2}\alpha Q - i\beta P_1$ is an isomorphism from $C_D^{0,\alpha}$ onto $C_T^{0,\alpha}$. Then $L$ is an isomorphism from $C_D^{0,\alpha} \times C_D^{0,\alpha}$ onto $C_D^{0,\alpha} \times C_T^{0,\alpha}$, provided that $k_1 + k_2 \neq 0$.

LEMMA 3.3. *Assume that $\hat{k}$ is not an eigenvalue of the boundary value problem* $\operatorname{curl}\mathbf{E} - i\hat{k}\mathbf{H} = 0$, $\operatorname{curl}\mathbf{H} + i\hat{k}\mathbf{E} = 0$ *in* $\Omega$, $\mathbf{n}\times\mathbf{E} = 0$ *on* $\partial\Omega$, *which therefore admits only the trivial solution* $\mathbf{E} \equiv 0$ *in* $\Omega$. *Furthermore, let* $\alpha \neq 0$ *and let* $\hat{k}, \beta$ *satisfy* $\operatorname{Re}(\hat{k}\beta/\alpha) > 0$ *on* $\partial\Omega$. *Let* $\widehat{\Phi}(\mathbf{x},\mathbf{y}) = e^{i\hat{k}|\mathbf{x}-\mathbf{y}|}/(4\pi|\mathbf{x}-\mathbf{y}|)$ *and let* $\widehat{P}, \widehat{M}$ *be the operators* $P_2, M_2$, *where* $\Phi_2$ *is replaced by* $\widehat{\Phi}$. *Then* $\frac{1}{2}\alpha Q - i\beta\widehat{P} + \alpha Q\widehat{M}$ *is an isomorphism from* $C_D^{0,\alpha}$ *onto* $C_T^{0,\alpha}$.

*Proof.* (i) *Injectivity.* Let $\mathbf{b} \in C_D^{0,\alpha}(\partial\Omega)$ with $\frac{1}{2}\alpha\mathbf{n}\times\mathbf{b} - i\beta\widehat{P}\mathbf{b} + \alpha\mathbf{n}\times\widehat{M}\mathbf{b} = 0$ on $\partial\Omega$. Set $\mathbf{E}(\mathbf{x}) = \operatorname{curl}^2\int_{\partial\Omega}\mathbf{b}(\mathbf{y})\widehat{\Phi}(\mathbf{x},\mathbf{y})\,ds(\mathbf{y})$, $\mathbf{H} = (1/i\hat{k})\operatorname{curl}\mathbf{E}$ in $\mathbf{R}^3\setminus\partial\Omega$. Then $\mathbf{E},\mathbf{H} \in C^1(\mathbf{R}^3\setminus\partial\Omega)\cap C(\overline{\Omega})\cap C(\mathbf{R}^3\setminus\Omega)$, and, from the jump conditions, we have

$$\frac{\alpha}{\hat{k}}\mathbf{n}\times(\mathbf{n}\times\mathbf{H})|_+ - \beta\mathbf{n}\times\mathbf{E}|_+ = \frac{1}{i}\left(\frac{\alpha}{2}\mathbf{n}\times\mathbf{b} + \alpha\mathbf{n}\times\widehat{M}\mathbf{b} - i\beta\widehat{P}\mathbf{b}\right) = 0.$$

Thus $(\mathbf{E},\mathbf{H})$ solves a homogeneous exterior impedance problem with boundary condition

$$\mathbf{n}\times(\mathbf{n}\times\mathbf{H})|_+ - \Psi\,\mathbf{n}\times\mathbf{E}|_+ = 0 \quad \text{on } \partial\Omega,$$

where $\Psi = \hat{k}\beta/\alpha$ on $\partial\Omega$. Since $\operatorname{Re}\Psi > 0$, the uniqueness result in [3] implies that $\mathbf{E} \equiv \mathbf{H} \equiv 0$ in $\mathbf{R}^3\setminus\Omega$. Since $\mathbf{n}\times\mathbf{E}|_- = \mathbf{n}\times\mathbf{E}|_- - \mathbf{n}\times\mathbf{E}|_+ = 0$ and we have the assumption that $\hat{k}$ is not a Maxwell eigenvalue in $\Omega$, we also have that $\mathbf{E} \equiv \mathbf{H} \equiv 0$ in $\Omega$, and thus, from the jump conditions for $\mathbf{n}\times\mathbf{H}$, we conclude that $\mathbf{b} = i/\hat{k}\,(\mathbf{n}\times\mathbf{H}|_+ - \mathbf{n}\times\mathbf{H}|_-) = 0$ on $\partial\Omega$, which shows injectivity.

(ii) *Surjectivity.* Let $\mathbf{c} \in C_T^{0,\alpha}(\partial\Omega)$ and $\mathbf{E},\mathbf{H} \in C^1(\mathbf{R}^3\setminus\overline{\Omega})\cap C(\mathbf{R}^3\setminus\Omega)$ be the solution to the exterior impedance problem

$$\begin{aligned} &\operatorname{curl}\mathbf{E} - i\hat{k}\mathbf{H} = 0, \qquad \operatorname{curl}\mathbf{H} + i\hat{k}\mathbf{E} = 0 \quad \text{in } \mathbf{R}^3\setminus\overline{\Omega},\\ &\frac{\alpha}{\hat{k}}\mathbf{n}\times(\mathbf{n}\times\mathbf{H})|_+ - \lambda\mathbf{n}\times\mathbf{E}|_+ = \frac{1}{i}\mathbf{c} \quad \text{on } \partial\Omega, \end{aligned} \tag{3.14}$$

which exists and is unique, as is proved in [3]. Indeed, the proof in [3], together with the standard estimates in [4, §2.6], show that $\mathbf{E}$ and $\mathbf{H}$ are Hölder continuous in $\mathbf{R}^3\setminus\Omega$. Furthermore, let $\mathbf{E},\mathbf{H} \in C^1(\Omega)\cap C(\overline{\Omega})$ be the unique solution of the interior Maxwell problem

$$\begin{aligned} &\operatorname{curl}\mathbf{E} - i\hat{k}\mathbf{H} = 0, \qquad \operatorname{curl}\mathbf{H} + i\hat{k}\mathbf{E} = 0 \quad \text{in } \Omega,\\ &\mathbf{n}\times\mathbf{E}|_- = \mathbf{n}\times\mathbf{E}|_+ \quad \text{on } \partial\Omega \end{aligned} \tag{3.15}$$

and set

$$\mathbf{b} := \frac{i}{\hat{k}}\,(\mathbf{n}\times\mathbf{H}_+ - \mathbf{n}\times\mathbf{H}|_-) \quad \text{on } \partial\Omega. \tag{3.16}$$

Then $\mathbf{b} \in C_D^{0,\alpha}(\partial\Omega)$, since $\operatorname{curl}\mathbf{H} = -i\hat{k}\mathbf{E} \in C^{0,\alpha}(\overline{\Omega})\cap C^{0,\alpha}(\mathbf{R}^3\setminus\Omega)$. Moreover,

$$\mathbf{E}(\mathbf{x}) = \operatorname{curl}^2\int_{\partial\Omega}\mathbf{b}(\mathbf{y})\hat{\Phi}(\mathbf{x},\mathbf{y})\,ds(\mathbf{y}), \qquad \mathbf{x} \in \mathbf{R}^3\setminus\partial\Omega, \tag{3.17}$$

$$\mathbf{H} = \frac{1}{i\hat{k}}\operatorname{curl}\mathbf{E} \quad \text{in } \mathbf{R}^3\setminus\partial\Omega. \tag{3.18}$$

In fact, if $\mathbf{E}_1$ and $\mathbf{H}_1$ denote the right-hand sides of (3.17) and (3.18), respectively, then $\mathbf{E}_1$ and $\mathbf{H}_1$ satisfy (3.15) and (3.16). By the above uniqueness argument, (3.17) and (3.18) are satisfied. The second equation of (3.14) then reads

$$\frac{1}{i}\mathbf{c} = \frac{1}{i}\left(\frac{\alpha}{2}\mathbf{n}\times\mathbf{b} + \alpha\mathbf{n}\times\widehat{M}\mathbf{b} - i\lambda\widehat{P}\mathbf{b}\right) \quad \text{on } \partial\Omega,$$

which proves surjectivity. This shows that $L$ is bijective. Thus $L^{-1}$ exists and is bounded according to the open mapping principle.

*Remark.* For $\lambda \equiv 0$, it is known (cf. [4]) that $\frac{1}{2}Q + Q\widehat{M}$ is an isomorphism from $C_D^{0,\alpha}$ onto $\{\mathbf{c} \in C_T^{0,\alpha} : \operatorname{Div}(\mathbf{n}\times\mathbf{c}) \in C^{0,\alpha}(\partial\Omega)\}$. This shows that the assumption $\operatorname{Re}(\hat{k}\lambda/\alpha) > 0$ is necessary.

We now apply this result to system (3.13), which can be written in the form

$$\tilde{L}\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix} + \tilde{K}\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix} = \begin{pmatrix}0\\ \mathbf{d}\end{pmatrix}, \tag{3.19}$$

where

$$\tilde{L} = L + \begin{pmatrix}0 & 0\\ 0 & \beta(P_1 - \widehat{P}) + \frac{\alpha}{i}Q\widehat{M}\end{pmatrix},$$

$$\tilde{K} = K - \begin{pmatrix}0 & 0\\ 0 & \beta(P_1 - \widehat{P}) + \frac{\alpha}{i}Q\widehat{M}\end{pmatrix}.$$

Lemma 3.3 establishes that $\tilde{L}$ is an isomorphism from $C_D^{0,\alpha}(\partial\Omega)\times C_D^{0,\alpha}(\partial\Omega)$ onto $C_D^{0,\alpha}(\partial\Omega)\times C_T^{0,\alpha}(\partial\Omega)$. The operator $\tilde{K}$ is clearly compact between these spaces. Hence (3.19) is equivalent to

$$\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix} + \tilde{L}^{-1}\tilde{K}\begin{pmatrix}\mathbf{a}\\ \mathbf{b}\end{pmatrix} = \tilde{L}^{-1}\begin{pmatrix}0\\ \mathbf{d}\end{pmatrix}, \tag{3.20}$$

which is a Fredholm equation of the second kind in $C_D^{0,\alpha}(\partial\Omega)\times C_D^{0,\alpha}(\partial\Omega)$.

To show uniqueness of solutions of (3.20), or equivalently of (3.13), we assume that the boundary value problem (3.1)–(3.4) itself has at most one solution (which holds, e.g., under assumption (3.6)).

Let $\binom{\mathbf{a}}{\mathbf{b}} \in C_D^{0,\alpha}(\partial\Omega)\times C_D^{0,\alpha}(\partial\Omega)$ be a solution of (3.13) for $\mathbf{d} = 0$ on $\partial\Omega$ and define

$$\mathbf{E}_j(\mathbf{x}) = k_j \operatorname{curl}\int_{\partial\Omega}\mathbf{a}(\mathbf{y})\Phi_j(\mathbf{x},\mathbf{y})\,ds(\mathbf{y}) + \operatorname{curl}^2\int_{\partial\Omega}\mathbf{b}(\mathbf{y})\Phi_j(\mathbf{x},\mathbf{y})\,ds(\mathbf{y}) \quad \text{in } \mathbf{R}^3\setminus\partial\Omega$$

and

$$\mathbf{H}_j(\mathbf{x}) = \frac{1}{ik_j}\operatorname{curl}\mathbf{E}_j(\mathbf{x}) \quad \text{in } \mathbf{R}^3\setminus\partial\Omega \ (j = 1,2).$$

Use of standard potential-theoretic arguments and the jump conditions then lead to the following theorem.

THEOREM 3.4. *Let $k_1$, $k_2$, $\mu_1$, $\mu_2 \in \mathbf{C}\setminus\{0\}$ with $\operatorname{Im} k_j \geq 0$ $(j = 1,2)$, $\mu_1 k_1 + \mu_2 k_2 \neq 0$, $k_1 + k_2 \neq 0$ and $\beta \in C^{0,\alpha}(\partial\Omega)$ with*

$$\operatorname{Re}\left(\frac{\beta(\mathbf{x})}{\mu_1 k_1 + \mu_2 k_2}\hat{k}\right) > 0 \tag{3.21}$$

*for all $x \in \partial\Omega$ and some $\hat{k}$ with $\operatorname{Im}\hat{k} \geq 0$. If the boundary value problem* (3.1)–(3.4) *has at most one solution, it has exactly one solution.*

*Remarks.* (a) From the boundary integral equations (3.10), (3.11), we see that the more general problem admits a unique solution for all $\mathbf{c} \in C_D^{0,\alpha}$, $\mathbf{d} \in C_T^{0,\alpha}$ (under the same assumptions of Theorem 3.4), as follows:

$$\begin{aligned}
&\operatorname{curl}\mathbf{E} - ik\mathbf{H} = 0, \qquad \operatorname{curl}\mathbf{H} + ik\mathbf{E} = 0 \quad \text{in } \mathbf{R}^3 \setminus \partial\Omega,\\
&\mathbf{n}\times\mathbf{E}|_+ - \mathbf{n}\times\mathbf{E}|_- = \mathbf{c} \ \text{ on } \partial\Omega,\\
&\mu_2\mathbf{n}\times(\mathbf{n}\times\mathbf{H})|_+ - \mu_1\mathbf{n}\times(\mathbf{n}\times\mathbf{H})|_- = \beta\mathbf{n}\times\mathbf{E} + \mathbf{d} \quad \text{on } \partial\Omega,\\
&\frac{\mathbf{x}}{|\mathbf{x}|}\times\mathbf{H}(\mathbf{x}) + \mathbf{E}(\mathbf{x}) = o\left(\frac{1}{|\mathbf{x}|}\right), \qquad |\mathbf{x}| \to \infty.
\end{aligned}$$

We see that, for $\lambda \equiv 0$, the assertion of the theorem cannot be valid for all $\mathbf{d} \in C_T^{0,\alpha}$, but at most for those $\mathbf{d} \in C_T^{0,\alpha}$ with $\mathbf{n}\times\mathbf{d} \in C_D^{0,\alpha}$ (since $\operatorname{Div}(\mathbf{n}\times\mathbf{H}) = \mathbf{n}\cdot\operatorname{curl}\mathbf{H} = -ik\mathbf{n}\cdot\mathbf{E} \in C^{0,\alpha}(\partial\Omega)$). This transmission problem has been considered by Wilde [23]. The limiting behavior for $\lambda \to 0$ is discussed in [8].

(b) In the physical situation described in §2, the various parameters are related by

$$\mu_1 k_1 + \mu_2 k_2 = k_1^2 + k_2^2 = i\omega\mu_0(\sigma_n + \sigma_a) \neq 0, \qquad k_1 + k_2 \neq 0$$

and

$$\frac{\beta(\mathbf{x})}{\mu_1 k_1 + \mu_2 k_2} = -i\frac{\tau(\mathbf{x})}{\sigma_n + \sigma_a}.$$

For $\hat{k} = i$, assumption (3.21) is satisfied.

4. Denseness of far field patterns. It is well known (cf. [4]) that the Silver–Müller radiation condition (3.4) implies the asymptotic behavior

$$\begin{aligned}
\mathbf{E}^s(\mathbf{x}) &= \frac{\exp(ik_2|\mathbf{x}|)}{|\mathbf{x}|}\left[\mathbf{E}_\infty(\hat{\mathbf{x}}) + \mathcal{O}\left(|\mathbf{x}|^{-1}\right)\right],\\
\mathbf{H}^s(\mathbf{x}) &= \frac{\exp(ik_2|\mathbf{x}|)}{|\mathbf{x}|}\left[\mathbf{H}_\infty(\hat{\mathbf{x}}) + \mathcal{O}\left(|\mathbf{x}|^{-1}\right)\right] \quad \text{as } |\mathbf{x}| \to \infty
\end{aligned}$$

uniformly in $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}| \in S^2$, with the properties

$$\mathbf{H}_\infty = \hat{\mathbf{x}}\times\mathbf{E}_\infty, \qquad \hat{\mathbf{x}}\cdot\mathbf{E}_\infty = \hat{\mathbf{x}}\cdot\mathbf{H}_\infty = 0 \quad \text{on } S^2.$$

The fields $\mathbf{E}_\infty, \mathbf{H}_\infty : S^2 \to \mathbf{C}^3$ are known as the far field patterns of $\mathbf{E}$, $\mathbf{H}$. The components of $\mathbf{E}_\infty$, $\mathbf{H}_\infty$ are analytic functions on $S^2$.

Moreover, the Stratton–Chu [18] representation theorem (cf. also [4])

$$\mathbf{E}^s(\mathbf{x}) = \operatorname{curl}\int_{\partial\Omega}\mathbf{e}(\mathbf{y})\Phi_2(\mathbf{x},\mathbf{y})ds(\mathbf{y}) - \frac{1}{ik_2}\operatorname{curl}^2\int_{\partial\Omega}\mathbf{h}(\mathbf{y})\Phi_2(\mathbf{x},\mathbf{y})ds(\mathbf{y}),$$

$\mathbf{x} \in \mathbf{R}^3 \setminus \overline{\Omega}$, with $\mathbf{e} := \mathbf{n}\times\mathbf{E}^s$, $\mathbf{h} := \mathbf{n}\times\mathbf{H}^s$, and the asymptotic form

$$\Phi_2(\mathbf{x},\mathbf{y}) = \frac{\exp(ik_2|\mathbf{x}|)}{|\mathbf{x}|}\left[e^{-ik_2\hat{\mathbf{x}}\cdot\mathbf{y}} + \mathcal{O}\left(|\mathbf{x}|^{-1}\right)\right] \quad \text{for } |\mathbf{x}| \to \infty,$$

uniformly in $\hat{\mathbf{x}} \in S^2$ and $\mathbf{y} \in \partial\Omega$, imply the representation

$$\mathbf{E}_\infty(\hat{\mathbf{x}}) = ik_2 \left[ \hat{\mathbf{x}} \times \int_{\partial\Omega} \mathbf{e}(\mathbf{y}) e^{-ik_2 \hat{\mathbf{x}}\cdot\mathbf{y}} ds(\mathbf{y}) + \hat{\mathbf{x}} \times \int_{\partial\Omega} (\mathbf{h}(\mathbf{y}) \times \hat{\mathbf{x}}) e^{-ik_2 \hat{\mathbf{x}}\cdot\mathbf{y}} ds(\mathbf{y}) \right], \tag{4.1}$$

for $\hat{\mathbf{x}} \in S^2$. This representation holds for any solution $(\mathbf{E}^s, \mathbf{H}^s)$ of Maxwell's equations (3.1) in $\mathbf{R}^3 \setminus \overline{\Omega}$.

Now we take the special case of the conductive boundary value problem (3.1)–(3.4) and denote the corresponding solution and far field pattern by $\mathbf{E}(\mathbf{x}, \hat{\mathbf{d}}, \mathbf{p})$ and $\mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{p})$, respectively, explicitly indicating the dependence on the direction $\hat{\mathbf{d}}$ and polarisation $\mathbf{p}$ of the incident plane wave.

Then we can prove the following reciprocity principle.

THEOREM 4.1. *Let $\beta$, $\mu_1$, $\mu_2$ be related in such a way that the conductive boundary value problem (3.1)–(3.4) admits a unique solution for every incident plane wave (see, e.g., the last remark of the preceding section). Then*

$$\mathbf{q} \cdot \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{p}) = \mathbf{p} \cdot \mathbf{E}_\infty(-\hat{\mathbf{d}}, -\hat{\mathbf{x}}, \mathbf{q})$$

*for all $\hat{\mathbf{x}}, \hat{\mathbf{d}} \in S^2$, $\mathbf{p}, \mathbf{q} \in \mathbf{C}^3$ with $\mathbf{p} \cdot \hat{\mathbf{d}} = 0$, $\mathbf{q} \cdot \hat{\mathbf{x}} = 0$.*

*Proof.* Let $\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{p}, \mathbf{q}$ be as indicated. Then $\mathbf{E}(\mathbf{x}, \hat{\mathbf{d}}, \mathbf{p}) = \mathbf{E}^s(\mathbf{x}, \hat{\mathbf{d}}, \mathbf{p}) + \mathbf{E}^i(\mathbf{x}, \hat{\mathbf{d}}, \mathbf{p})$ with $\mathbf{E}^i(\mathbf{x}, \hat{\mathbf{d}}, \mathbf{p}) = \mathbf{p} \exp(ik_2 \hat{\mathbf{d}} \cdot \mathbf{x})$, and we can decompose $\mathbf{H}$ in the same way with $\mathbf{H}^i(\mathbf{x}, \hat{\mathbf{d}}, \mathbf{p}) = (\hat{\mathbf{d}} \times \mathbf{p}) \exp(ik_2 \hat{\mathbf{d}} \cdot \mathbf{x})$. Then (4.1) implies that

$$\begin{aligned} \frac{1}{ik_2} \mathbf{q} \cdot \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{p}) = &\int_{\partial\Omega} \left( \mathbf{n} \times \mathbf{E}^s(\mathbf{y}, \hat{\mathbf{d}}, \mathbf{p}) \right) \cdot \mathbf{H}^i(\mathbf{y}, -\hat{\mathbf{x}}, \mathbf{q}) \, ds(\mathbf{y}) \\ &+ \int_{\partial\Omega} \left( \mathbf{n} \times \mathbf{H}^s(\mathbf{y}, \hat{\mathbf{d}}, \mathbf{p}) \right)\Big|_- \cdot \mathbf{E}^i(\mathbf{y}, -\hat{\mathbf{x}}, \mathbf{q}) \, ds(\mathbf{y}), \end{aligned}$$

since $\mathbf{q} \cdot (\hat{\mathbf{x}} \times \mathbf{e}) = \mathbf{e} \cdot (-\hat{\mathbf{x}} \times \mathbf{q})$ and $\mathbf{q} \cdot \hat{\mathbf{x}} = 0$. If $(\mathbf{E}^s, \mathbf{H}^s)$ and $(\mathbf{F}^s, \mathbf{G}^s)$ satisfy Maxwell's equations and the radiation condition, Green's formula in $\mathbf{R}^3 \setminus \Omega$ yields

$$\int_{\partial\Omega} [(\mathbf{n} \times \mathbf{E}^s) \cdot \mathbf{G}^s + (\mathbf{n} \times \mathbf{H}^s) \cdot \mathbf{F}^s] \, ds = 0.$$

Hence, if we choose $\mathbf{F}^s = \mathbf{E}^s(\cdot, -\hat{\mathbf{x}}, \mathbf{q})$,

$$\begin{aligned} \frac{1}{ik_2} \mathbf{q} \cdot \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{p}) = &\int_{\partial\Omega} \left( \mathbf{n} \times \mathbf{E}^s(\cdot, \hat{\mathbf{d}}, \mathbf{p}) \right) \cdot \mathbf{H}(\cdot, -\hat{\mathbf{x}}, \mathbf{q})|_+ \, ds \\ &+ \int_{\partial\Omega} \left( \mathbf{n} \times \mathbf{H}^s(\cdot, \hat{\mathbf{d}}, \mathbf{p}) \right)\Big|_+ \cdot \mathbf{E}(\cdot, -\hat{\mathbf{x}}, \mathbf{q}) \, ds. \end{aligned}$$

Now $\mathbf{n} \times (\mathbf{E}|_+ \times \mathbf{n}) = \mathbf{n} \times (\mathbf{E}|_- \times \mathbf{n})$ on $\partial\Omega$ and $\mu_2 \mathbf{n} \times (\mathbf{H}|_+ \times \mathbf{n}) - \mu_1 \mathbf{n} \times (\mathbf{H}|_- \times \mathbf{n}) = -\beta \mathbf{n} \times \mathbf{E}|_-$ on $\partial\Omega$; thus

$$\begin{aligned} &\frac{\mu_2}{ik_2} \mathbf{q} \cdot \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{p}) \\ &= \int_{\partial\Omega} \left[ \mathbf{n} \times \mathbf{E}(\cdot, \hat{\mathbf{d}}, \mathbf{p}) - \mathbf{n} \times \mathbf{E}^i(\cdot, \hat{\mathbf{d}}, \mathbf{p}) \right] \cdot \left[ \mu_1 \mathbf{H}(\cdot, -\hat{\mathbf{x}}, \mathbf{q})|_- - \beta \mathbf{n} \times \mathbf{E}(\cdot, -\hat{\mathbf{x}}, \mathbf{q}) \right] ds \\ &\quad + \int_{\partial\Omega} \mathbf{E}(\cdot, -\hat{\mathbf{x}}, \mathbf{q}) \cdot \left[ \mu_1 \mathbf{n} \times \mathbf{H}(\cdot, \hat{\mathbf{d}}, \mathbf{p})\Big|_- - \beta \mathbf{n} \times (\mathbf{n} \times \mathbf{E}(\cdot, \hat{\mathbf{d}}, \mathbf{p})) \right. \\ &\quad \left. - \mu_2 \mathbf{n} \times \mathbf{H}^i(\cdot, \hat{\mathbf{d}}, \mathbf{p}) \right] ds. \end{aligned}$$

Now we use Green's formula in $\Omega$ to find that

$$\int_{\partial\Omega}\left[\Big(\mathbf{n}\times\mathbf{E}(\cdot,\hat{\mathbf{d}},\mathbf{p})\Big)\cdot\mathbf{H}(\cdot,-\hat{\mathbf{x}},\mathbf{q})|_{-}+\mathbf{E}(\cdot,-\hat{\mathbf{x}},\mathbf{q})\cdot\Big(\mathbf{n}\times\mathbf{H}(\cdot,\hat{\mathbf{d}},\mathbf{p})\Big|_{-}\Big)\right]ds=0,$$

from which it follows that

$$\begin{aligned}
&\frac{\mu_2}{ik_2}\mathbf{q}\cdot\mathbf{E}_\infty(\hat{\mathbf{x}},\hat{\mathbf{d}},\mathbf{p})\\
&=-\int_{\partial\Omega}\Big(\mathbf{n}\times\mathbf{E}^{\mathrm{i}}(\cdot,\hat{\mathbf{d}},\mathbf{p})\Big)\cdot\mu_2\mathbf{H}(\cdot,-\hat{\mathbf{x}},\mathbf{q})\Big|_{+}ds\\
&\quad-\int_{\partial\Omega}\mathbf{E}(\cdot,-\hat{\mathbf{x}},\mathbf{q})\cdot\mu_2\Big(\mathbf{n}\times\mathbf{H}^{\mathrm{i}}(\cdot,\hat{\mathbf{d}},\mathbf{p})\Big)ds\\
&=\mu_2\int_{\partial\Omega}\big(\mathbf{n}\times\mathbf{H}(\cdot,-\hat{\mathbf{x}},\mathbf{q})|_{+}\big)\cdot\mathbf{p}e^{ik_2\hat{\mathbf{d}}\cdot}+(\mathbf{n}\times\mathbf{E}(\cdot,-\hat{\mathbf{x}},\mathbf{q}))\cdot(\hat{\mathbf{d}}\times\mathbf{p})e^{ik_2\hat{\mathbf{d}}\cdot}ds\\
&=\frac{\mu_2}{ik_2}\mathbf{p}\cdot\mathbf{E}_\infty(-\hat{\mathbf{d}},-\hat{\mathbf{x}},\mathbf{q}),
\end{aligned}$$

which proves the theorem.

Now we define

$$V:=\operatorname{span}\left\{(\hat{\mathbf{d}}\times\mathbf{p})e^{ik_2\hat{\mathbf{d}}\cdot}\Big|_{\partial\Omega}:\hat{\mathbf{d}}\in S^2,\ \mathbf{p}\in\mathbf{C}^3\right\},$$

the space of all possible linear combinations of plane incident fields on $\partial\Omega$, and denote by $\mathcal{F}_V$ the space of corresponding far field patterns. Define $\mathcal{E}$ by

$$(\mathbf{E}_1,\mathbf{E}_2)\in\mathcal{E}\text{ if and only if there exist }\mathbf{H}_1,\mathbf{H}_2\text{ with}$$

$$\operatorname{curl}\mathbf{E}_j=ik_j\mathbf{H}_j,\qquad\operatorname{curl}\mathbf{H}_j=-ik_j\mathbf{E}_j\quad\text{in }\Omega,\quad j=1,2,$$

$$\mathbf{n}\times\mathbf{E}_2-\mathbf{n}\times\mathbf{E}_1=0\quad\text{on }\partial\Omega, \tag{4.2}$$

$$\mu_2\mathbf{n}\times(\mathbf{n}\times\mathbf{H}_2)-\mu_1\mathbf{n}\times(\mathbf{n}\times\mathbf{H}_1)=\beta\mathbf{n}\times\mathbf{E}_1\quad\text{on }\partial\Omega.$$

Then we may state and prove the following result.

THEOREM 4.2. *Assume that, for the given values of $\beta$, $\mu_1$, $\mu_2$, the conductive boundary value problem (3.1)–(3.4) admits a unique solution for every incident plane wave. Then the orthogonal complement $\mathcal{F}_V^{\perp}$ of $\mathcal{F}_V$ in $L_T^2(S^2)$ is given by*

$$\mathcal{F}_V^{\perp}=\left\{\overline{\mathbf{h}}\in L_T^2(S^2):\ \textit{there exists }\mathbf{E}_1\textit{ such that }(\mathbf{E}_1,\mathbf{E}_{\mathbf{h}})\in\mathcal{E}\right\}, \tag{4.3}$$

*where*

$$\mathbf{E}_{\mathbf{h}}(\mathbf{x})=\int_{S^2}\mathbf{h}(\hat{\mathbf{d}})e^{-ik_2\hat{\mathbf{d}}\cdot\mathbf{x}}ds(\hat{\mathbf{d}}),\quad\mathbf{x}\in\Omega, \tag{4.4}$$

*denotes the so-called Herglotz field with kernel* $\mathbf{h}$.

*Proof.* Suppose first that $\mathbf{h}\in L_T^2(S^2)$ with $\overline{\mathbf{h}}\perp\mathcal{F}_V$, i.e.,

$$\int_{S^2}\mathbf{n}(\hat{\mathbf{x}})\cdot\mathbf{E}_\infty(\hat{\mathbf{x}},\hat{\mathbf{d}},\hat{\mathbf{d}}\times\mathbf{p})ds(\hat{\mathbf{x}})=0\qquad\text{for all }\hat{\mathbf{d}}\in S^2,\quad\mathbf{p}\in\mathbf{C}^3.$$

Then, by the reciprocity principle (since $\mathbf{h}(\hat{\mathbf{x}}) \cdot \hat{\mathbf{x}} = \hat{\mathbf{d}} \cdot (\hat{\mathbf{d}} \times \mathbf{p}) = 0$),

$$\int_{S^2} (\hat{\mathbf{d}} \times \mathbf{p}) \cdot \mathbf{E}_\infty(-\hat{\mathbf{d}}, -\hat{\mathbf{x}}, \mathbf{h}(\hat{\mathbf{x}})) ds(\hat{\mathbf{x}}) = 0 \qquad \text{for all } \hat{\mathbf{d}} \in S^2, \quad \mathbf{p} \in \mathbf{C}^3.$$

Interchanging the roles of $-\hat{\mathbf{x}}$ and $\hat{\mathbf{d}}$, we see that

$$(\hat{\mathbf{x}} \times \mathbf{p}) \cdot \int_{S^2} \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{h}(-\hat{\mathbf{d}})) ds(\hat{\mathbf{d}}) = 0 \qquad \text{for all } \hat{\mathbf{x}} \in S^2, \quad \mathbf{p} \in \mathbf{C}^3.$$

The function

$$\int_{S^2} \mathbf{E}_\infty(\cdot, \hat{\mathbf{d}}, \mathbf{h}(-\hat{\mathbf{d}})) ds(\hat{\mathbf{d}})$$

is the far field pattern $\mathbf{F}_\infty$ of the solution $(\mathbf{F}, \mathbf{G})$ of the conductive boundary value problem (3.1)–(3.4) with incident field

$$\mathbf{E}^i(\mathbf{x}) = \int_{S^2} \mathbf{h}(-\hat{\mathbf{d}}) e^{ik_2 \hat{\mathbf{d}} \cdot \mathbf{x}} ds(\hat{\mathbf{d}}) = \int_{S^2} \mathbf{h}(\hat{\mathbf{d}}) e^{-ik_2 \hat{\mathbf{d}} \cdot \mathbf{x}} ds(\hat{\mathbf{d}}) = \mathbf{E}_\mathbf{h}(\mathbf{x}).$$

Since $(\hat{\mathbf{x}} \times \mathbf{p}) \cdot \mathbf{F}_\infty(\hat{\mathbf{x}}) = 0$ for all $\hat{\mathbf{x}} \in S^2$, $\mathbf{p} \in \mathbf{C}^3$, the far field $\mathbf{F}_\infty$ vanishes on $S^2$; thus $\mathbf{F}^s$ and $\mathbf{G}^s$ vanish in $\mathbf{R}^3 \setminus \Omega$.

We define $\mathbf{E}_1 := \mathbf{F}$, $\mathbf{H}_1 := \mathbf{G}$ and $\mathbf{E}_2 := \mathbf{E}^i = \mathbf{E}_\mathbf{h}$, $\mathbf{H}_2 := (1/ik_2)\operatorname{curl} \mathbf{E}_\mathbf{h}$ in $\Omega$. Then $(\mathbf{E}_1, \mathbf{E}_2)$ solves (4.2). Hence $\mathcal{F}^\perp$ is contained in the set given in the right-hand side of (4.3). The opposite inclusion follows from reversing the preceeding arguments.

System (4.2) can be considered as an eigenvalue problem for the two parameters $k_1$, $k_2$. By essentially the same arguments as in §3, we can establish the Fredholm alternative for this system: If (4.2) admits only the trivial solution $\mathbf{E}_j = 0$ in $\Omega$ for $j = 1, 2$, then the inhomogeneous form of (4.2) has a unique solution for every inhomogeneity. In this case, Theorem 4.2 states that $\mathcal{F}_V$ is dense in $L^2_T(S^2)$. Instead of considering $V$, we then look at the far field patterns generated by the space

$$A := \operatorname{span}\left\{ (\hat{\mathbf{d}} \times \mathbf{p}) e^{ik_2 \hat{\mathbf{d}} \cdot} \Big|_{\partial\Omega} - (\hat{\mathbf{d}}_1 \times \mathbf{p}) e^{ik_2 \hat{\mathbf{d}}_1 \cdot} \Big|_{\partial\Omega} : \hat{\mathbf{d}} \in S^2, \mathbf{p} \in \mathbf{C}^3 \right\}$$

for a fixed direction $\hat{\mathbf{d}}_1 \in S^2$. Let $\mathcal{F}_A$ be the space of the corresponding far field patterns and $\mathbf{h} \in L^2_T(S^2)$ with $\overline{\mathbf{h}} \in \mathcal{F}_A^\perp$; i.e.,

$$\int_{S^2} \mathbf{h}(\hat{\mathbf{x}}) \cdot \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \hat{\mathbf{d}} \times \mathbf{p})\, ds(\hat{\mathbf{x}}) = \int_{S^2} \mathbf{h}(\hat{\mathbf{x}}) \cdot \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}_1, \hat{\mathbf{d}}_1 \times \mathbf{p})\, ds(\hat{\mathbf{x}})$$

for all $\hat{\mathbf{d}} \in S^2$, $\mathbf{p} \in \mathbf{C}^3$. Using the reciprocity principle as before and interchanging $\hat{\mathbf{d}}$ and $\hat{\mathbf{x}}$ yields

$$(\hat{\mathbf{x}} \times \mathbf{p}) \cdot \int_{S^2} \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{h}(-\hat{\mathbf{d}}))\, ds(\hat{\mathbf{d}}) = -ik_2 \mathbf{p} \cdot \mathbf{c} \qquad \text{for all } \hat{\mathbf{x}} \in S^2, \quad \mathbf{p} \in \mathbf{C}^3,$$

with $\mathbf{c} := (1/ik_2)\hat{\mathbf{d}}_1 \times \int_{S^2} \mathbf{E}_\infty(-\hat{\mathbf{d}}_1, -\hat{\mathbf{x}}, \mathbf{h}(\hat{\mathbf{x}}))\, ds(\hat{\mathbf{x}})$, i.e.,

$$\mathbf{p} \cdot \left[ \hat{\mathbf{x}} \times \int_{S^2} \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{h}(\hat{\mathbf{d}}))\, ds(\hat{\mathbf{d}}) - ik_2 \mathbf{c} \right] = 0 \qquad \text{for all } \hat{\mathbf{x}} \in S^2, \quad \mathbf{p} \in \mathbf{C}^3.$$

From this, we conclude that

$$\hat{\mathbf{x}} \times \int_{S^2} \mathbf{E}_\infty(\hat{\mathbf{x}}, \hat{\mathbf{d}}, \mathbf{h}(-\hat{\mathbf{d}}))\, ds(\hat{\mathbf{d}}) = ik_2\mathbf{c} \quad \text{for all } \hat{\mathbf{x}} \in S^2.$$

As before, the function

$$\int_{S^2} \mathbf{E}_\infty(\cdot, \hat{\mathbf{d}}, \mathbf{h}(-\hat{\mathbf{d}}))\, ds(\hat{\mathbf{d}})$$

is the far field pattern $\mathbf{F}_\infty$ of the solution $(\mathbf{F}, \mathbf{G})$ of the conductive boundary value problem (3.1)–(3.4) with incident field $\mathbf{E}^i = \mathbf{E}_\mathbf{h}$. Since the far field pattern $\mathbf{F}_\infty$ is a tangential field, we conclude that $\mathbf{F}_\infty(\hat{\mathbf{x}}) = -ik_2\hat{\mathbf{x}} \times \mathbf{c}$ for all $\hat{\mathbf{x}} \in S^2$. Thus $\mathbf{F}_\infty$ coincides with the far field pattern of $-\operatorname{curl}[\mathbf{c}(\exp(ik_2|\mathbf{x}|)/|\mathbf{x}|)]$. From the uniqueness of the far field patterns, the corresponding fields must coincide; thus $\mathbf{F}^s(\mathbf{x}) = -\operatorname{curl}[\mathbf{c}(\exp(ik_2|\mathbf{x}|)/|\mathbf{x}|)]$ for $\mathbf{x}$ on $\mathbf{R}^3 \setminus \Omega$. Again, we define $\mathbf{E}_1 := \mathbf{F}$, $\mathbf{H}_1 := \mathbf{G}$, $\mathbf{E}_2 := \mathbf{E}^i = \mathbf{E}_\mathbf{h}$, $\mathbf{H}_2 := (1/ik_2)\operatorname{curl}\mathbf{E}_\mathbf{h}$ in $\Omega$. Then $(\mathbf{E}_1, \mathbf{E}_2)$ solves

$$\operatorname{curl}\mathbf{E}_j = ik_j\mathbf{H}_j, \qquad \operatorname{curl}\mathbf{H}_j = -ik_j\mathbf{E}_j \quad \text{in } \Omega,\ j = 1, 2,$$

$$\mathbf{n} \times \mathbf{E}_2 - \mathbf{n} \times \mathbf{E}_1 = \mathbf{n} \times \operatorname{curl}\left[\mathbf{c}\frac{\exp(ik_2|\mathbf{x}|)}{|\mathbf{x}|}\right] \quad \text{on } \partial\Omega, \tag{4.5}$$

$$\mu_2\mathbf{n} \times (\mathbf{n} \times \mathbf{H}_2) - \mu_1\mathbf{n} \times (\mathbf{n} \times \mathbf{H}_1) = \beta\mathbf{n} \times \mathbf{E}_1 + \frac{\mu_2}{ik_2}\mathbf{n} \times \left(\mathbf{n} \times \operatorname{curl}^2\left[\mathbf{c}\frac{\exp(ik_2|\mathbf{x}|)}{|\mathbf{x}|}\right]\right)$$

on $\partial\Omega$. This argument establishes the following theorem.

THEOREM 4.3. *Let $\beta$, $\mu_1$, $\mu_2$ be such that the conductive boundary value problem (3.1)–(3.4) admits a unique solution for every incident wave $\mathbf{E}^i$. Then the orthogonal complement of $\mathcal{F}_A^\perp$ of $\mathcal{F}_A$ in $L_T^2(S^2)$ is given by*

$$\mathcal{F}_A^\perp = \{\overline{\mathbf{h}} \in L_T^2(S^2) : \text{ there exists } \mathbf{E}_1 \text{ and } \mathbf{c} \in \mathbf{C}^3 \text{ such that } (\mathbf{E}_1, \mathbf{E}_\mathbf{h}) \text{ solves (4.5)}\}.$$

## REFERENCES


[1] T. S. ANGELL, R. E. KLEINMAN, AND F. HETTLICH, *The resistive and conductive problems for the exterior Helmholtz equation*, SIAM J. Appl. Math., 50 (1990), pp. 1607–1622.

[2] A. P. CALDERÒN, *The multipole expansion of radiation fields*, J. Rational Mech. Anal., 3 (1954), pp. 523–537.

[3] D. L. COLTON AND R. KRESS, *The impedance boundary value problem for the time harmonic Maxwell equations*, Math. Meth. Appl. Sci., 3 (1981), pp. 475–487.

[4] ———, *Integral Equations in Scattering Theory*, John Wiley, New York, 1983.

[5] G. A. GRAY AND R. E. KLEINMAN, *The integral equation method in electromagnetic scattering*, J. Math. Anal. Appl., 107 (1985), pp. 455–477.

[6] R. F. HARRINGTON AND J. R. MAUTZ, *An impedance sheet approximation for thin dielectric shells*, IEEE Trans. Antennás and Propagation, AP-23 (1975), pp. 531–534.

[7] F. HETTLICH, *Die Integralgleichungsmethode bei Streuung an Körpern mit einer dünnen Schicht*, Diplomarbeit, Göttingen, 1989.

[8] A. KIRSCH, *Two singular perturbation problems for Maxwell's equations*, preprint.

[9] W. KNAUFF AND R. KRESS, *On the exterior boundary-value problem for the time-harmonic Maxwell equations*, J. Math. Anal. Appl., 72 (1979), pp. 215–235.

[10] R. KRESS, *On the boundary operator in electromagnetic scattering*, Proc. Roy. Soc. Edinburgh, 103A (1986), pp. 91–98.

[11] E. MARX, *A single integral equation for wave scattering*, J. Math. Phys., 23 (1982), pp. 1057–1065.

[12] ———, *Integral equations for scattering by a dielectric*, IEEE Trans. Antennas and Propagation, AP-32 (1984), pp. 166–172.

[13] A. W. MAUE, *Zur Formulierung eines allgemeiner Beugung problems durch eine Integralgleichung*, Z. Phys., 126 (1949), pp. 601–618.

[14] C. MÜLLER, *Über die Beugung elektromagnetischer Schwingungen an endlichen homogenen Körpern*, Math. Ann., 123 (1951), pp. 345–378.

[15] W. K. SAUNDERS, *On solutions of Maxwell's equations in an exterior region*, Proc. Nat. Acad. Sci. U.S.A., 38 (1952), pp. 342–348.

[16] U. SCHMUCKER, *Interpretation of induction anomalies above nonuniform surface layers*, Geophys., 36 (1971), pp. 156–165.

[17] T. B. A. SENIOR, *A note on the impedance boundary condition*, Canad. J. Phys., 40 (1962), pp. 663–665.

[18] J. A. STRATTON AND L. J. CHU, *Diffraction theory of electromagnetic waves*, Phys. Rev., 56 (1939), pp. 99–107.

[19] G. VASSEUR AND P. WEIDELT, *Bimodal electromagnetic induction in nonuniform thin sheets with an application to the northern Pyrenean induction anomaly*, Geophys. J. Roy. Astronom. Soc., 51 (1977), pp. 669–690.

[20] P. WERNER, *On the exterior boundary value problem of perfect reflection for stationary electromagnetic wave fields*, J. Math. Anal. Appl., 7 (1963), pp. 348–396.

[21] ———, *On an integral equation in electromagnetic diffraction theory*, J. Math. Anal. Appl., 14 (1966), pp. 445–462.

[22] H. WEYL, *Die natürlichen Randwertaufgaben im Aussenraum für Strahlungsfelder beliebigen Dimension und beliebiger Ranges*, Math. Zeit., 56 (1952), p. 105.

[23] P. WILDE, *Transmission problems for the vector Helmholtz equation*, Proc. Roy. Soc. Edinburgh, 105A (1987), pp. 61–76.

Mathematical Methods in the Applied Sciences, Vol. 15, 647–660 (1992)
MOS subject classification: 78 A 50, 49 B 22

# Multicriteria Optimization in Antenna Design

T. S. Angell

*Department of Mathematical Sciences, University of Delaware, Newark, Delaware 19716, U.S.A.*

and

A. Kirsch

*Institute of Applied Mathematics, University of Erlangen and Nurenberg, D-8520 Erlangen, Germany*

Communicated by D. S. Jones

**We formulate certain problems in the optimal design of radiating structures such as multicriteria optimization problems. We review the basic background of such problems, prove the existence of Pareto optimal points, and give necessary conditions. We apply the latter to the numerical computation of optimal surface currents for the problem of simultaneously optimizing both the quality factor and the signal-to-noise ratio of a conformal antenna.**

## 1. Introduction

In our earlier paper [1] we summarized a coherent approach to the problem of optimal antenna design. By formulating various measures of performance as real-valued functionals defined in the appropriate function spaces, we can systematically use the methods of functional analysis and optimization theory not only to study the existence and properties of optimal solutions but also to develop computational procedures for the numerical approximation of these solutions in concrete cases. Thus in [1] we provided analyses, including computational results, for two specific cases: the maximization of power radiated into a given sector (or sectors) of the far-field region and the problem of maximizing the signal-to-noise ratio.

It has long been recognized [11], [8] (esp. ch. 8) that the narrow focusing of the main beam of an antenna has the concomitant effect of increasing the near-field power. Not only may one wish to focus the main beam, but also to minimize the power stored in the near-field region. Thus we see a typical problem that arises in antenna design, namely the problem of dealing with several, possibly conflicting, goals.

The approach used most often in such situations is illustrated by the two examples in [1]. In the first case we introduced a constraint on the power available to the

0170–4214/92/090647–14$12.00
*Received 13 August 1991*

antenna by considering surface currents that are bounded in some appropriate norm: in the second, we required the so-called quality factor, that is the ratio of input power to the far-field power (both measured, again, relative to the appropriate function-space norms), to be bounded.

In this paper we wish to suggest another approach to such antenna design problems, namely the approach of multicriteria optimization. While well known in other applied fields these techniques have not been applied, to our knowledge, to problems of antenna design. The subject of multicriteria optimization has been most thoroughly developed in the literature of mathematical economics and is most often associated there with the names of Walras and Pareto who introduced the basic notions in the late 1890s. The interested reader may consult the review article of Stadler [9] for the historical background and the article of Dauer and Stadler [3] for more recent developments. Applications to problems in mechanical engineering are described in [10], which has an extensive bibliography.

We dedicate the following section to an outline of the necessary background material including the general conditions ensuring the existence of Pareto points and necessary conditions in the form of a multiplier rule. Section 3 contains a statement of the optimization problem and the proof of existence of an optimal solution. The final section contains a numerical example.

## 2. Notation and basic theorems

We recall that a convex cone $\Lambda$ in a linear space $Z$ is a convex set with the property that

$$x \in \Lambda,\ \alpha \geqslant 0 \text{ implies } \alpha x \in \Lambda. \tag{2.1}$$

Note that, in particular, $0 \in \Lambda$. Such a cone defines a partial order, $<_\Lambda$, on $Z$ according to

$$x <_\Lambda y \text{ provided } y - x \in \Lambda. \tag{2.2}$$

In order to ensure that the relation is not only reflexive and transitive, but also antisymmetric we require, further, that the cone be *pointed*, that is that $\Lambda \cap (-\Lambda) = \{0\}$. In this case $x <_\Lambda y$ and $y <_\Lambda x$ implies that $y - x = 0$.

**Example 2.1.** *The most common example is that for which* $Z = \mathbb{R}^n$ *and*

$$\Lambda = \{x = (x^1, \ldots, x^n) | x^i \geqslant 0,\ i = 1, 2, \ldots, n\}.$$

*Then* $x <_\Lambda y$ *if and only if* $x^i \leqslant y^i$ *for all* $i = 1, 2, \ldots, n$ *where the latter inequality involves the usual ordering in* $\mathbb{R}$.

**Example 2.2.** *Let* $Z = SL_n(\mathbb{R}^n)$, *the set of symmetric* $n \times n$ *matrices, and set*

$$\Lambda = \{A \in SL_n(\mathbb{R}^n) | (Ax, x) \geqslant 0 \text{ for all } x \in \mathbb{R}^n\}.$$

*Then* $\Lambda$ *is a convex, pointed cone.*

In problems of vector optimization we are interested in *minimal elements* relative to a given order cone.

**Definition 2.3.** *Let* $S \neq \varnothing$ *be a subset of an ordered vector space* $Z$. *Then* $x_0 \in S$ *is a minimal element of* $S$ *provided* $x \in S$ *and* $x <_\Lambda x_0$ *implies* $x = x_0$.

**Example 2.4.** *Let $Z = \mathbb{R}^2$ with $\Lambda = \{(x, y) | x \geqslant 0, y \geqslant 0\}$ as order cone. Let*

$$S = \{(x, y) | x \geqslant 0, y \geqslant 0, xy \geqslant 1\}.$$

*Then all the points of the set*

$$\{(x, y) \in S | xy = 1\}$$

*are minimal.*

The general multicriteria optimization problem can now be formulated as follows: given a linear space $X$ and an ordered linear space $Z$, let $U \subset X$ and suppose $g: U \to Z$. Find $u_0 \in U$ such that $g(u_0)$ is a minimal element of $g(U)$.

**Definition 2.5.** *A point $u_0 \in U$ is said to be Pareto optimal relative to the vector-valued function $g$ provided $g(u_0)$ is minimal with respect to $g(U)$.*

The term Pareto optimal is chosen here for historical reasons. Other terms have been frequently used including 'non-inferior solutions', 'non-dominated solutions', and 'efficient solutions'. Some of these terms may be more informative in that they better suggest the property that characterizes Pareto points, namely that we cannot lower one of the component values by moving from that point without strictly increasing at least one of the other components of the criterion vector.

In general, Pareto points are not unique as we can see in the following simple example.

**Example 2.6.** *Consider the cone $\Lambda \subset \mathbb{R}^2$ of Example 2.4. Then the Pareto set i.e. the set of Pareto optimal points for the function*

$$g(x, y) = \begin{pmatrix} x \\ y \end{pmatrix}$$

*defined on $S = \{(x, y) | x \geqslant 0, y \geqslant 0, xy \geqslant 1\}$, is just the set*

$$\{(x, y) \in S | xy = 1\}.$$

We remark that, in general, it is not true that there exists some point that will minimize all the components of the vector criterion simultaneously, nor is it necessarily true that standard scalar optimization methods can be used to find the Pareto set. In particular, it is not generally the case that the minimization of *one* criterion subject to inequality constraints on the others will yield a Pareto point.

In order to develop conditions guaranteeing that a point is Pareto optimal, we need to introduce the concept of a *polar* cone. To this end, let $Z$ be a Banach space with dual $Z^*$. Thus $Z^*$ is the set of all continuous linear maps $\lambda: Z \to \mathbb{R}$. Denote the action of an element $\lambda \in Z^*$ on $z \in Z$ by $\langle \lambda, z \rangle$. For an arbitrary set $E \subset Z$ we have

**Definition 2.7.** *Let $E \subset Z$. The polar of the set $E$ is defined to be*

$$E^p := \{\lambda \in Z^* | \langle \lambda, z \rangle \leqslant 0 \text{ for all } z \in E\}. \tag{2.3}$$

Note that, by the linearity and continuity of $\lambda$, $E^p$ is a closed convex cone in $Z^*$ regardless of the nature of the set $E$. We shall refer to $E^p$ as the *polar cone* of $E$. Related to the cone $E^p$ is the set $E^-$ given by

$$E^- := \{\lambda \in E^p | \langle \lambda, z \rangle < 0 \text{ for all } z \in E, z \neq 0\}. \tag{2.4}$$

It is clear that $E^- \cup \{0\}$ is likewise a convex cone in $Z^*$. We shall call it the *strict polar*

*cone of E*. It is a standard result [4] that the inclusion

$$\operatorname{int}(\Lambda^p) \subset \Lambda^- \tag{2.5}$$

is valid provided $Z$ is reflexive and $\Lambda$ is a non-trivial closed convex cone in $Z$.

It is now a simple matter to establish a basic sufficient condition for a point to be minimal.

**Proposition 2.8.** *Let $Z$ be a real Banach space ordered by a non-trivial convex cone $\Lambda$ and let $\bar{\lambda} \in \Lambda^-$. If $S \subset Z$ and $y \in S$ is such that*

$$\langle \bar{\lambda}, z \rangle \leqslant \langle \bar{\lambda}, y \rangle \text{ for all } z \in S \tag{2.6}$$

*then $y$ is a minimal point of $S$.*

*Proof.* Let $z_0 \in S$, $z_0 \neq y$, and suppose $z_0 <_\Lambda y$. Then $(y - z_0) \in \Lambda \setminus \{0\}$. Hence $\langle \bar{\lambda}, y - z_0 \rangle < 0$, which contradicts (2.6).

We can now prove a theorem guaranteeing the existence of Pareto points under conditions of wide applicability. We need only recall the following definition.

**Definition 2.9.** *A map $g: X \to Z$, $X$, $Z$ Banach spaces, is called completely continuous provided $g$ maps weakly convergent sequences into norm convergent sequences.*

We can now state the following existence theorem.

**Theorem 2.10.** *Let $X$ be a Banach space and $Z$ an ordered reflexive Banach space with a non-trivial closed convex order cone $\Lambda$. Suppose that $\operatorname{int}(\Lambda^p) \neq \varnothing$. Then if $U$ is a closed bounded convex subset of $X$ and*

$$g: U \to Z$$

*is completely continuous, then $g(U)$ has a minimal point and $U$ contains a Pareto point.*

*Proof.* By Mazur's theorem [12] $U$ is weakly compact and hence, by the complete continuity of $g$, $g(U)$ is compact in $Z$. If we can show that $g(U)$ contains a minimal point $\bar{z}$, then any $\bar{u} \in U$ such that $g(\bar{u}) = \bar{z}$ is a Pareto point.

Let $\bar{\lambda} \in \operatorname{int}(\Lambda^p)$. Then $\bar{\lambda} \neq 0$ for, if $\bar{\lambda} = 0$, then $\Lambda^p$ could contain a ball and hence would coincide with $Z^*$. By reflexivity $(\Lambda^p)^p = \Lambda$ so that $\Lambda = \{0\}$, which would contradict the assumption that $\Lambda$ is non-trivial.

Now, for the given $\bar{\lambda}$ consider the map

$$z \to \langle \bar{\lambda}, z \rangle, \quad z \in g(U)$$

of $g(U) \to \mathbb{R}$. By continuity and compactness this map has a maximum on $g(U)$, say at $\bar{z} \in g(U)$. Then

$$\langle \bar{\lambda}, \bar{z} \rangle \geqslant \langle \bar{\lambda}, z \rangle \text{ for all } z \in g(U).$$

Thus $\bar{z}$ is a minimal point by Proposition 2.8.

The assumption of complete continuity of the map $g$ in Theorem 2.10 implies, in particular, that $g$ maps bounded sets into bounded sets. Thus Example 2.6 shows that Pareto points may well exist even if $g$ does not have this property. Indeed the set of Pareto points may itself be unbounded. That problems with unbounded sets $g(U)$ may arise in applications, we shall see below in section 3. In such cases it may be possible to show that the existence of Pareto points follows from the compactness of certain subsets of $g(U)$. In fact, we have the following result.

**Theorem 2.11.** *Let $X$ be a Banach space and $Z$ an ordered reflexive Banach space with a non-trivial closed convex order cone $\Lambda$. Suppose that $\operatorname{int}(\Lambda^p) \neq \varnothing$ and that $U \subset X$. Then if $g: U \to Z$ is such that, for some $z \in Z$,*

$$G_z = (z - \Lambda) \cap g(U) \tag{2.7}$$

*is non-empty and compact, then $U$ contains a Pareto point.*

*Proof.* Note that if $z_0$ is a minimal point for $G_z$ then $z_0$ is also minimal for $g(U)$. Indeed since $z_0$ is minimal for $G_z$,

$$(z_0 - \Lambda) \cap G_z = \{z_0\}. \tag{2.8}$$

We have to show that $(z_0 - \Lambda) \cap g(U) = \{z_0\}$. Thus let $x \in g(U) \cap (z_0 - \Lambda)$. Then $x \in z_0 - \Lambda$ and also $z_0 \in z - \Lambda$. By adding these inclusions we obtain $x \in z - \Lambda$, that is $x \in (z_0 - \Lambda) \cap G_z = \{z_0\}$. Thus $z_0$ is also minimal for $g(U)$. Now, if $G_z$ is compact in $Z$ then the proof of Theorem 2.10 shows that $G_z$ has a minimal element and the proof is complete.

We conclude this section with a necessary condition, in the form of a Lagrange multiplier rule, which will be suitable for the specific problem discussed in section 3 below. A more general statement, as well as the proof, can be found in the book of Kirsch, Warth and Werner [5].

**Theorem 2.12.** *Let $X$ and $Z$ be Banach spaces satisfying the hypotheses of Theorem 2.10 and suppose that $g: X \to Z$ is Frèchet differentiable while $h: X \to \mathbb{R}$ is continuously Frèchet differentiable. Let*

$$S = \{x \in X \mid h(x) = 0\}$$

*and suppose that $\tilde{x} \in S$ is a Pareto point for $g$. Then there exists a $\tilde{\lambda} \in -\Lambda^p$ and a $\mu \geqslant 0$ such that*

$$[\tilde{\lambda} g'(\tilde{x}) + \mu h'(\tilde{x})]x = 0 \quad \text{for all } x \in X.$$

*Moreover, if $h'(\tilde{x})$ is surjective then $\tilde{\lambda} \neq 0$.*

*Remark.* This theorem holds even for weak Pareto points $\tilde{x}$, that is for those $\tilde{x} \in S$ that satisfy

$$[g(\tilde{x}) - \operatorname{int} \Gamma] \cap g(S) = \varnothing$$

provided that $\operatorname{int} \Gamma \neq \varnothing$.

## 3. An optimization problem for antenna design

As described in our earlier paper [1], we consider as an antenna any radiating structure that supports a flow of charge, or surface current I, and which thereby produces an electromagnetic field in a homogeneous isotropic medium exterior to the structure. For definiteness, we consider here the case of a connected region $D \subset \mathbb{R}^3$, with non-empty interior, $D_-$, and $C^2$-boundary $S$. We shall use $D_+$ for the (connected) exterior domain $\mathbb{R}^3 \setminus (S \cup D_-)$, we shall denote points by their position vectors x and y, and we shall choose the origin of the coordinates to lie in $D_-$.

Assuming a harmonic time dependence, $e^{-i\omega t}$, the field, (E, H), produced by I is required to satisfy the time-harmonic form of the Maxwell equations

$$\nabla \times \mathbf{E} - ikZ_0\mathbf{H} = 0, \tag{3.1a}$$

$$\nabla \times \mathbf{H} + ikY_0\mathbf{E} = 0, \tag{3.1b}$$

where $Z_0 = (\mu_0/\varepsilon_0)^{1/2}$, $Y_0 = 1/Z_0$, $k = \omega(\varepsilon_0\mu_0)^{1/2}$ and $\varepsilon_0$, $\mu_0$ are the free-space permittivity and permeability, respectively. The quantities $Z_0$ and $Y_0$ are the free-space impedance and admittance.

We recall [2] that the fields E and H have the asymptotic representation

$$\mathbf{E}(\mathbf{x}) = \frac{e^{ikr}}{r}\mathbf{F}(\hat{\mathbf{x}}) + \mathcal{O}(1/r^2), \quad r \to \infty, \tag{3.2a}$$

and

$$\mathbf{H}(\mathbf{x}) = Y_0\frac{e^{ikr}}{r}\hat{\mathbf{x}} \times \mathbf{F}(\hat{\mathbf{x}}) + \mathcal{O}(1/r^2), \quad r \to \infty, \tag{3.2b}$$

where $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$ and $r = |\mathbf{x}|$.

The vector function F, which has no radial component, is called the radiation pattern. It is an analytic function defined on the unit ball $S^1$.

The problems that we summarized in [1] involve some numerical measure of performance, which is to be optimized by selecting the appropriate surface current, I, from some preassigned subset of admissible currents. The existence and uniqueness of a solution to (3.1) satisfying the boundary condition

$$Z_0\hat{n} \times \mathbf{H} = \mathbf{I} \text{ on } S \tag{3.3}$$

and the Silver–Müller radiation condition [2], p. 113 for every tangential field

$$\mathbf{I} \in \mathbf{L}_t^2(S) = \{\Psi \in \mathbf{L}^2(S): \hat{n} \cdot \Psi = 0 \text{ on } S\}$$

guarantee the existence of a mapping

$$\mathscr{K}: \mathbf{L}_t^2(S) \to \mathbf{L}_t^2(S^1), \tag{3.4}$$

which associates to each admissible current I the corresponding far-field pattern F. This map, which is not known explicitly except in certain special cases, is known to be compact and, by Corollary 4.10 of Colton and Kress [2], it is one-to-one.

In terms of this compact operator we can introduce several different measures of antenna performance. An extensive list appears in [1]. Here, we shall consider two examples of optimization problems based on such criteria: the first problem is one related to the problem of antenna synthesis, while the second (and more complicated) example involves the concept of signal-to-noise ratio.

The classical problem of antenna synthesis can be formulated as follows (see e.g., [8]). Given a desired far-field pattern $\mathbf{F}_0$, find the surface current I whose far field produces $\mathbf{F}_0$.

Stated in this way, the problem has no solution in general as $\mathbf{F}_0$ may not be an actual far field. In particular $\mathbf{F}_0$ may not be analytic. However, as shown in [2], the range of $\mathscr{K}$ is dense in $\mathbf{L}_t^2(S^1)$, and we usually formulate the problem as that of finding a best approximation to $\mathbf{F}_0$ measured in some suitable norm. For example, it is

common to consider the functional

$$D(\mathbf{I}) := \int_{S^1} |\mathscr{K}\mathbf{I}(\mathbf{x}) - \mathbf{F}_0(\mathbf{x})|^2 \, ds. \tag{3.5}$$

Good approximations in this sense can be realized only by producing unacceptable levels of the 'quality factor' given by

$$\mathscr{Q}(\mathbf{I}) := \|\mathbf{I}\|^2_{\mathbf{L}^2_t(S)} / \|\mathscr{K}\mathbf{I}\|^2_{\mathbf{L}^2_t(S^1)}, \tag{3.6}$$

which compares the power radiated into the far-field region with the power supplied to the antenna structure. Here we suggest that appropriate compromises can be studied by identifying the Pareto points for the vector criterion

$$\mathscr{Q}(\mathbf{I}) := \begin{pmatrix} D(\mathbf{I}) \\ -\|\mathscr{K}\mathbf{I}\|^2_{\mathbf{L}^2_t(S^1)} \end{pmatrix} \tag{3.7}$$

subject to the power constraint

$$\|\mathbf{I}\|_{\mathbf{L}^2_t(S)} \leqslant 1. \tag{3.8}$$

We are assured of the existence of Pareto points for this problem by the following result:

**Theorem 3.1.** *The map $\mathscr{Q}: \mathbf{L}^2_t(S) \to \mathbb{R}^2$ is completely continuous and hence Pareto points exist.*

*Proof.* Since the relatively compact sets in $\mathbb{R}^2$ are the bounded sets, it suffices to show that $\mathscr{Q}$ maps bounded sets into bounded sets. However, this follows immediately from the boundedness of the operator $\mathscr{K}$ and the fact that $S$ is a bounded surface of finite area.

As a second example we introduce the signal-to-noise ratio (SNR) defined by

$$SNR(\hat{\mathbf{x}}, \mathbf{I}) := \frac{|(\mathscr{K}\mathbf{I})(\hat{\mathbf{x}})|^2}{\int_{S^1} \omega(\hat{\mathbf{y}})^2 |\mathscr{K}\mathbf{I}(\hat{\mathbf{y}})|^2 \, ds}, \tag{3.9}$$

where $\omega \in \mathbf{L}^\infty(S^1)$ is non-zero on a set $T$ of positive measure. The denominator of (3.9) is a measure of how much the radiated field is corrupted by noise. For a fixed direction $\hat{\mathbf{x}}$ and constant $c > 0$, we can formulate an optimization problem as

maximize SNR(I)

subject to $\|\mathbf{I}\|^2_{\mathbf{L}^2_t(S)} \leqslant 1$ and $\mathscr{Q}(\mathbf{I}) \leqslant c$,

where $\mathscr{Q}$ is given by (3.6).

This problem was studied, for the acoustic case, by Kirsch and Wilde [6] and was formulated, as we have done here, for the full three-dimensional electromagnetic problem in [1]. These results generalize the problem for planar apertures that were first studied by Lo *et al.* [7]. Here we wish to consider, not this constrained problem, but the vector-valued problem with the criterion

$$\mathbf{V}(\mathbf{I}) := \begin{pmatrix} -\mathrm{SNR}(\mathbf{I}) \\ \mathscr{Q}(\mathbf{I}) \end{pmatrix} \tag{3.10}$$

subject to

$$\mathbf{I} \neq 0. \tag{3.11}$$

In order to prove that Pareto points for the problem (3.10), (3.11) exist we shall use Theorem 2.11.

Let $\mathscr{V}$ be the set of attainable points, that is

$$\mathscr{V} := \{\mathbf{V}(\mathbf{I}) : \mathbf{I} \neq 0\}. \tag{3.12}$$

Since we cannot show the closedness of $\mathscr{V}$, we extend this set, show the existence of the Pareto points of this extended set and prove afterwards that these Pareto points lie in fact in $\mathscr{V}$. We begin with the following result.

**Lemma 3.2.** *The set*

$$\mathscr{V}_0 := \left\{\mathbf{V}(\mathbf{I}) + \begin{pmatrix} 0 \\ r \end{pmatrix} : \mathbf{I} \neq 0, r \geqslant 0\right\}$$

*is closed and $\mathscr{V}_0 \cap (z - \Lambda)$ is compact for every $z \in \mathscr{V}_0$. Hence $\mathscr{V}_0$ has Pareto points.*

*Proof.* To show that $\mathscr{V}_0$ is closed let $\{\mathbf{I}_n\} \subset \mathbf{L}_t^2(S)$, $\{r_n\} \subset \mathbb{R}$ with $r_n \geqslant 0$, $\mathbf{I}_n \neq 0$ and

$$\mathbf{V}(\mathbf{I}_n) + \begin{pmatrix} 0 \\ r_n \end{pmatrix} \to \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{R}^2.$$

Since $\mathbf{V}$ is scale invariant, that is for any scalar $z \in \mathbb{C}\backslash\{0\}$, $\mathbf{V}(z\mathbf{I}) = \mathbf{V}(\mathbf{I})$, we can assume that $\|\mathbf{I}_n\|_{\mathbf{L}_t^2(S)} = 1$ and thus $\{\mathbf{I}_n\}$ contains a weak-limit point. Without loss of generality we assume that $\mathbf{I}_n \to \mathbf{I}$ weakly in $\mathbf{L}_t^2(S)$ for some $\mathbf{I}$ with $\|\mathbf{I}\|_{\mathbf{L}_t^2(S)} \leqslant 1$. Since $\mathscr{K}$ is compact both as a map into $\mathbf{L}_t^2(S^1)$ and into $\mathbf{C}_t(S^1)$ we have that $\mathscr{K}\mathbf{I}_n \to \mathscr{K}\mathbf{I}$ in $\mathbf{L}_t^2(S^1)$ and $\mathbf{C}_t(S^1)$. We show that $\mathbf{I} \neq 0$. This follows from the convergence

$$\mathscr{Q}(\mathbf{I}_n) + r_n = \|\mathscr{K}\mathbf{I}_n\|_{\mathbf{L}_t^2(S^1)}^{-2} + r_n \to z_2$$

since $r_n \geqslant 0$ and $\|\mathscr{K}\mathbf{I}_n\|_{\mathbf{L}_t^2(S^1)} \to \|\mathscr{K}\mathbf{I}\|_{\mathbf{L}_t^2(S^1)}$. Furthermore

$$r_n - r := z_2 - \|\mathscr{K}\mathbf{I}\|_{\mathbf{L}_t^2(S^1)}^{-2} \geqslant 0.$$

Also we have that $\omega\mathscr{K}\mathbf{I}_n \to \omega\mathscr{K}\mathbf{I}$ in $\mathbf{L}_t^2(S^1)$ and $\omega\mathscr{K}\mathbf{I} \neq 0$ because of the analyticity of $\mathscr{K}\mathbf{I}$. Therefore $\mathrm{SNR}(\mathbf{I}_n) \to \mathrm{SNR}(\mathbf{I}) = -z_1$. Thus we have

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} -SNR(\mathbf{I}) \\ r + \|\mathscr{K}\mathbf{I}_n\|_{\mathbf{L}_t^2(S^1)}^{-2} \end{pmatrix} = \mathbf{V}(\mathbf{I}) + \begin{pmatrix} 0 \\ s \end{pmatrix} \text{ with } s = r + \frac{1 - \|\mathbf{I}\|_{\mathbf{L}_t^2(S)}^2}{\|\mathscr{K}\mathbf{I}\|_{\mathbf{L}_t^2(S^1)}^2} \geqslant 0.$$

This means that

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathscr{V}_0,$$

which proves that $\mathscr{V}_0$ is closed.

To show the compactness of $\mathscr{V}_0 \cap (z - \Gamma)$ we only have to prove that this set is bounded. Let

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \quad \text{and} \quad \mathbf{V}(\mathbf{I}_n) + \begin{pmatrix} 0 \\ r_n \end{pmatrix}$$

be a sequence in $\mathscr{V}_0 \cap (z - \Gamma)$. Again we can assume that $\|\mathbf{I}_n\|_{\mathbf{L}_t^2(S)} = 1$ and $r_n \geqslant 0$.

Certainly $0 \leqslant \mathcal{Q}(\mathbf{I}_n) + r_n \leqslant z_2$ and $\mathrm{SNR}(\mathbf{I}_n) \geqslant 0$. Again we conclude as in the first part of this proof that $(\mathbf{I}_n)$ contains a weak limit point $\mathbf{I}$ with $\|\mathbf{I}\|_{\mathbf{L}_t^2(S)} \leqslant 1$ and $\mathscr{K}\mathbf{I}_n \to \mathscr{K}\mathbf{I}$ in $\mathbf{L}_t^2(S^1)$ and $\mathbf{C}_t(S^1)$. From $\mathcal{Q}(\mathbf{I}_n) \leqslant z_2$ we conclude that $\|\mathscr{K}\mathbf{I}_n\|^2_{\mathbf{L}_t^2(S^1)} \geqslant 1 - z_2$ and thus $\mathscr{K}\mathbf{I} \neq 0$. This shows that $\mathrm{SNR}(\mathbf{I}_n) \to \mathrm{SNR}(\mathbf{I})$, that is

$$\mathbf{V}(\mathbf{I}_n) + \begin{pmatrix} 0 \\ r_n \end{pmatrix}$$

is bounded.

In the light of Theorem 2.11 we know that $\mathscr{V}_0$ contains Pareto points. This ends the proof.

It is now very easy to show that existence of the Pareto points of $\mathscr{V}$.

**Theorem 3.3.** *Let*

$$z^* \in \mathscr{V}_0 := \left\{ \mathbf{V}(\mathbf{I}) + \begin{pmatrix} 0 \\ r \end{pmatrix} \middle| \mathbf{I} \neq 0, r \geqslant 0 \right\}$$

*be a Pareto point of $\mathscr{V}_0$. Then $z^* \in \mathscr{V}$, that is $z^*$ is also a Pareto point of $\mathscr{V}$.*

*Proof.* Let

$$z^* = \mathbf{V}(\mathbf{I}^*) + \begin{pmatrix} 0 \\ r^* \end{pmatrix}, \quad \mathbf{I}^* \neq 0, r^* \geqslant 0.$$

Since $\mathbf{V}(\mathbf{I}^*) \in \mathscr{V} \subset \mathscr{V}_0$ and $\mathbf{V}(\mathbf{I}^*) <_A z^*$ we conclude from the minimality of $z^*$ that $z^* = \mathbf{V}(\mathbf{I}^*) \in \mathscr{V}$. This completes the proof.

Now we shall apply Theorem 2.12 to the optimization problem (3.10), (3.11) and will use the resulting equations to compute the set of all 'critical points', which, as in the case of a single-cost functional, contains the set of Pareto points. Note that, for this particular example we shall not need the full force of that theorem since we have no explicit equality constraints. Certainly other situations will arise in applications where it will be useful to be able to handle such constraints and we may bring the full force of Theorem 2.12 to bear in such cases.

The Fréchet derivatives of SNR and $\mathcal{Q}$ at $\mathbf{I}_0 \in \mathbf{L}_t^2(S)$ are (here and in the following we write $\|\cdot\|$ for the $L^2$- norm, either on $S$ or on $S^1$):

$$\mathrm{SNR}'(\mathbf{I}_0, \mathbf{I}) = \frac{2}{\|\omega\mathscr{K}\mathbf{I}_0\|^4}[\|\omega\mathscr{K}\mathbf{I}_0\|^2 \mathrm{Re}(\mathscr{K}\mathbf{I}_0(\hat{\mathbf{x}}) \cdot \mathscr{K}\mathbf{I}(\hat{\mathbf{x}})) - |\mathscr{K}\mathbf{I}_0(\hat{\mathbf{x}})|^2 \mathrm{Re}\langle \omega\mathscr{K}\mathbf{I}_0, \omega\mathscr{K}\mathbf{I}\rangle],$$

$$\mathcal{Q}'(\mathbf{I}_0, \mathbf{I}) = \frac{2}{\|\mathscr{K}\mathbf{I}_0\|^4}[\|\mathscr{K}\mathbf{I}_0\|^2 \mathrm{Re}\langle \mathbf{I}_0, \mathbf{I}\rangle - \|\mathbf{I}_0\|^2 \mathrm{Re}\langle \mathscr{K}\mathbf{I}_0, \mathscr{K}\mathbf{I}\rangle].$$

We now assume for simplicity that $D$ is an infinite cylinder in the $z$-direction whose constant cross section we also denote by $D \subset \mathbb{R}^2$ (with boundary $S$). Furthermore let $\mathbf{I} = I\hat{\mathbf{z}}$ with a scalar function $I \in L^2(S)$. Finally, let $\mathscr{K}$ be given in the explicit form

$$\mathscr{K}I(\hat{\mathbf{x}}) := \sigma \int_S I(y) e^{-ik\hat{\mathbf{x}}\cdot y}\, ds(y) = \sigma\langle I, e^{ik\hat{\mathbf{x}}\cdot}\rangle_{L^2(S)} \quad \text{with } \sigma = \frac{i}{2}\sqrt{\frac{k}{2\pi}}\, e^{-i\pi/4}.$$

Let $I_0$ be a Pareto point. Application of Theorem 2.12 yields the existence of $\lambda, \mu \geq 0$ with $\lambda + \mu > 0$ and $-\lambda \mathrm{SNR}'(I_0, I) + \mu \mathcal{Q}'(I_0, I) = 0$ for all $I \in L^2(S)$, that is

$$-\frac{\lambda}{\|\omega \mathcal{K} I_0\|^4}\left[\|\omega \mathcal{K} I_0\|^2 (\mathcal{K} I_0(\hat{\mathbf{x}}) \bar{\sigma} e^{ik\hat{\mathbf{x}}\cdot} - |\mathcal{K} I_0(\hat{\mathbf{x}})|^2 \mathcal{K}^*(\omega^2 \mathcal{K} I_0)\right]$$

$$+\frac{\mu}{\|\mathcal{K} I_0\|^4}\left[\|\mathcal{K} I_0\|^2 I_0 - \|I_0\|^2 \mathcal{K}^* \mathcal{K} I_0\right] = 0,$$

or

$$\lambda \frac{|\mathcal{K} I_0(\hat{\mathbf{x}})|^2}{\|\omega \mathcal{K} I_0\|^4} \mathcal{K}^*(\omega^2 \mathcal{K} I_0) + \frac{\mu}{\|\mathcal{K} I_0\|^2} I_0 - \frac{\mu}{\|\mathcal{K} I_0\|^2} Q_0 \mathcal{K}^* \mathcal{K} I_0$$

$$= \frac{\lambda \mathcal{K} I_0(\hat{\mathbf{x}}) \bar{\sigma}}{\|\omega \mathcal{K} I_0\|^2} e^{ik\hat{\mathbf{x}}\cdot}$$

where $Q_0 = \|I_0\|^2 / \|\mathcal{K} I_0\|^2$.

Now we distinguish between two cases. If $\lambda \mathcal{K} I_0(\hat{\mathbf{x}}) = 0$ then $\mathcal{K}^* \mathcal{K} I_0 = (1/Q_0) I_0$, that is $1/Q_0$ is an eigenvalue of $\mathcal{K}^* \mathcal{K}$ with eigenfunction $I_0$.

If $\lambda > 0$ and $\mathcal{K} I_0(\hat{\mathbf{x}}) \neq 0$ we set

$$\tilde{\mu} := \frac{\mu}{\lambda \|\mathcal{K} I_0\|^2} \frac{\|\omega \mathcal{K} I_0\|^4}{|\mathcal{K} I_0(\hat{\mathbf{x}})|^2} \quad \text{and} \quad I_1 := \frac{\overline{\mathcal{K} I_0(\hat{\mathbf{x}})}}{\|\omega \mathcal{K} I_0\|^2} I_0.$$

Then $I_1$ is also Pareto optimal and $\mathcal{K} I_1(\hat{\mathbf{x}}) = \|\omega \mathcal{K} I_1\|^2$, thus $\mathrm{SNR}(I_1) = \|\omega \mathcal{K} I_1\|^2$, and

$$\mathcal{K}^*(\omega^2 \mathcal{K} I_1) + \tilde{\mu} I_1 - \tilde{\mu} Q_0 \mathcal{K}^* \mathcal{K} I_1 = \bar{\sigma} e^{ik\hat{\mathbf{x}}\cdot} \text{ on } S. \tag{3.13}$$

Therefore we see that if $I_1$ is a Pareto point of (3.10), (3.11), that is normalized so that $\mathcal{K} I_1(\hat{\mathbf{x}}) = \|\omega \mathcal{K} I_1\|^2$ then there exists $\tilde{\mu} \geq 0$ with (3.13) where $Q_0 = \|I_1\|^2 / \|\mathcal{K} I_1\|^2$. Solutions of (3.13) are called *critical Pareto points* of the problem (3.10), (3.11).

If, on the other hand, $I_1$ solves (3.13) for some $Q_0$ and $\tilde{\mu} > 0$ then

$$\mathcal{K} I_1(\hat{\mathbf{x}}) = \langle I_1, \bar{\sigma} e^{ik\hat{\mathbf{x}}\cdot} \rangle = \|\omega \mathcal{K} I_1\|^2 + \tilde{\mu} \|I_1\|^2 - \tilde{\mu} Q_0 \|\mathcal{K} I_1\|^2,$$

that is

$$\mathcal{K} I_1(\hat{\mathbf{x}}) - \|\omega \mathcal{K} I_1\|^2 = \tilde{\mu}[\|I_1\|^2 - Q_0 \|\mathcal{K} I_1\|^2].$$

Hence

$$\mathcal{K} I_1(\hat{\mathbf{x}}) = \|\omega \mathcal{K} I_1\|^2 \text{ is equivalent to } \|I_1\|^2 = Q_0 \|\mathcal{K} I_1\|^2. \tag{3.14}$$

Equations (3.13) and (3.14) describe a one-parameter family of critical points that contains the set of Pareto points or even the weak Pareto points.

We want to illustrate this approach via the necessary optimality conditions with a numerical example (cf. Kirsch and Wilde [6] for the related example where $\mathcal{Q}$ is fixed and SNR is to be maximized).

We consider the case where $D$ is the unit disc in $\mathbb{R}^2$ and $\omega$ is the characteristic function of a portion of the unit circle, for example,

$$\omega(t) = \begin{cases} 1 & \text{if } t_1 \leq t \leq t_2, \\ 0 & \text{otherwise.} \end{cases}$$

For this choice we can compute $\mathcal{K}$ explicitly: let

$$I(s) := I_1(s) = \sum_{j=-\infty}^{\infty} x_j e^{ijs}.$$

Then

$$\mathcal{K}I(t) = \sigma \sum_{j=-\infty}^{\infty} x_j \int_0^{2\pi} e^{ijs} e^{-ik\cos(t-s)}\, ds$$

$$= 2\pi\sigma \sum_{j=-\infty}^{\infty} x_j (-i)^j J_j(k) e^{ijt},$$

where we have used the fact that

$$e^{-ik\cos\tau} = \sum_{n=-\infty}^{\infty} (-i)^n J_n(k) e^{in\tau}.$$

Here, $J_n$ denotes the Bessel function of order $n$.

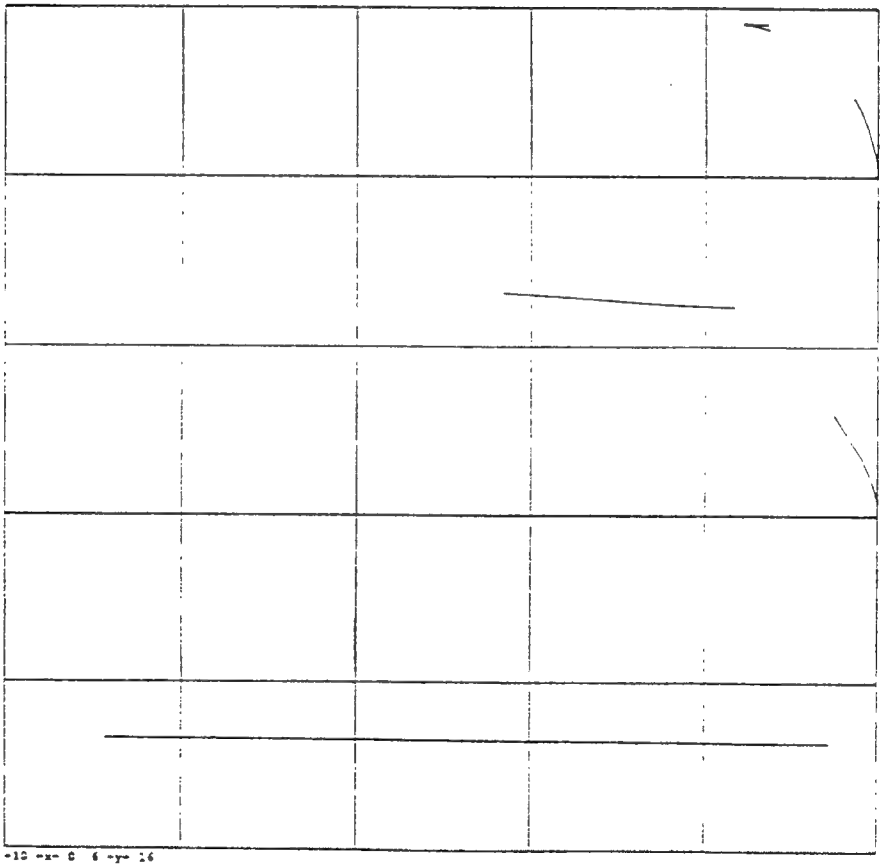


Fig. 1. Solution branches

Thus we see that $\mathscr{K}$ is an infinite diagonal matrix with elements $2\pi\sigma(-\mathrm{i})^j J_j(k)$. Likewise $\mathscr{K}^*\mathscr{K}$ is diagonal with elements $d_j := 4\pi^2|\sigma|^2 J_j(k)^2$ and

$$\mathscr{K}^*(\omega^2 \mathscr{K} I)(t) = 2\pi|\sigma|^2 \sum_{j=-\infty}^{\infty} x_j(-\mathrm{i})^j J_j(k) \int_{t_1}^{t_2} \mathrm{e}^{\mathrm{i}js}\mathrm{e}^{\mathrm{i}k\cos(t-s)}\,\mathrm{d}s$$

$$= \sum_{l=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} a_{lj} x_j \mathrm{e}^{\mathrm{i}lt}$$

where the coefficients $a_{lj}$ are defined by

$$a_{lj} = 2\pi|\sigma|^2(-\mathrm{i})^j J_j(k)\mathrm{i}^l J_l(k) \int_{t_1}^{t_2} \mathrm{e}^{\mathrm{i}(j-l)s}\,\mathrm{d}s$$

$$= \begin{cases} 2\pi|\sigma|^2 \mathrm{i}^{l-j} J_j(k) J_l(k) \dfrac{1}{\mathrm{i}(j-l)} (\mathrm{e}^{\mathrm{i}(j-l)t_2} - \mathrm{e}^{\mathrm{i}(j-l)t_1}), & \text{if } j \neq l, \\ 2\pi|\sigma|^2 J_j(k)^2 (t_2 - t_1), & \text{if } j = l. \end{cases}$$

We project equation (3.13) onto the finite-dimensional space

$$X_n = \operatorname{span}\{\mathrm{e}^{\mathrm{i}jt} : |j| \leqslant n\}.$$

Table 1. Solutions of Equation (3.15)

| $\tilde{\mu}$ | $Q_0$ | SNR | $Q_0$ | SNR | $Q_0$ | SNR | $Q_0$ | SNR |
|---|---|---|---|---|---|---|---|---|
| 0.2 | 7.39 | 36.23 | 11.20 | 0.52 | 12.64 | 4.32 | | |
| 0.3 | 7.38 | 30.46 | 10.67 | 0.18 | 12.58 | 3.46 | | |
| 0.4 | 7.37 | 26.99 | 10.50 | 0.10 | 12.54 | 3.01 | | |
| 0.5 | 7.36 | 24.43 | 10.36 | 0.06 | 12.52 | 2.75 | | |
| 0.6 | 7.36 | 22.32 | 10.29 | 0.04 | 12.51 | 2.57 | | |
| 0.7 | 7.36 | 20.51 | 10.25 | 0.03 | 12.50 | 2.44 | 14.94 | 0.28 |
| 0.8 | 7.35 | 18.92 | 10.21 | 0.03 | 12.50 | 2.35 | 14.78 | 0.20 |
| 0.9 | 7.35 | 17.50 | 10.18 | 0.02 | 12.49 | 2.27 | 14.67 | 0.15 |
| 1.0 | 7.35 | 16.23 | 10.16 | 0.02 | 12.49 | 2.22 | 14.58 | 0.12 |
| 1.1 | 7.35 | 15.09 | | | | | | |
| 1.2 | 7.34 | 14.08 | | | | | | |
| 1.3 | 7.34 | 13.17 | | | | | | |
| 1.4 | 7.34 | 12.33 | | | | | | |
| 1.5 | 7.34 | 11.58 | | | | | | |
| 1.6 | 7.33 | 10.92 | | | | | | |
| 1.7 | 7.33 | 10.34 | | | | | | |
| 1.8 | 7.33 | 09.78 | | | | | | |
| 1.9 | 7.33 | 09.28 | | | | | | |
| 2.0 | 7.33 | 08.84 | 10.06 | 0.00 | 12.48 | 1.95 | 14.22 | 0.03 |
| 3.0 | 7.31 | 06.01 | 10.01 | 0.00 | 12.48 | 1.86 | 14.11 | 0.01 |
| 4.0 | 7.31 | 04.65 | 10.01 | 0.00 | 12.48 | 1.82 | 14.05 | 0.01 |
| 5.0 | 7.30 | 03.88 | 09.91 | 0.00 | 12.47 | 1.80 | 14.02 | 0.00 |
| 6.0 | 7.30 | 03.38 | 09.91 | 0.00 | 12.47 | 1.77 | 14.00 | 0.00 |
| 7.0 | 7.30 | 03.04 | 09.81 | 0.00 | 12.47 | 1.75 | 13.98 | 0.00 |
| 8.0 | 7.29 | 02.78 | 09.81 | 0.00 | 12.47 | 1.75 | 13.97 | 0.00 |
| 9.0 | 7.29 | 02.60 | 09.71 | 0.00 | 12.47 | 1.73 | 13.96 | 0.00 |
| 10.0 | 7.29 | 02.45 | 09.61 | 0.00 | 12.47 | 1.73 | 13.95 | 0.00 |

Thus equations (3.13) and (3.14) take the form

$$(A + \tilde{\mu}E - \tilde{\mu}Q_0 D)x = r, \quad \sum_{j=-n}^{n} (1 - Q_0 d_j)|x_j|^2 = 0, \tag{3.15}$$

where $E$ is the identity matrix,

$$D = \operatorname{diag}(d_j : |j| \leqslant n), \text{ with } d_j = 4\pi^2 |\sigma_1|^2 J_j(k)^2,$$

$$A = (a_{lj})_{l,\, j=-n,\ldots,n} \quad \text{and} \quad r_j = \bar{\sigma}_1 i^j J_j(k) e^{-ij\delta}$$

where $\hat{\mathbf{x}} = (\cos\delta, \sin\delta)$.

We have computed this example for the choice $k = 6$, $t_1 = 40°$, $t_2 = 140°$, $\delta = -90°$ and $n = 16$. For large ranges of $\tilde{\mu}$ (from 0.2 to 20) we computed all the zeros of the function

$$\varphi(Q_0)_{\tilde{\mu}} := \sum_{j=-n}^{n} (1 - Q_0 d_j)|x_j|^2,$$

where $x$ solves the first equation of (3.15) for $Q_0$, by a simple bisection method. It turns

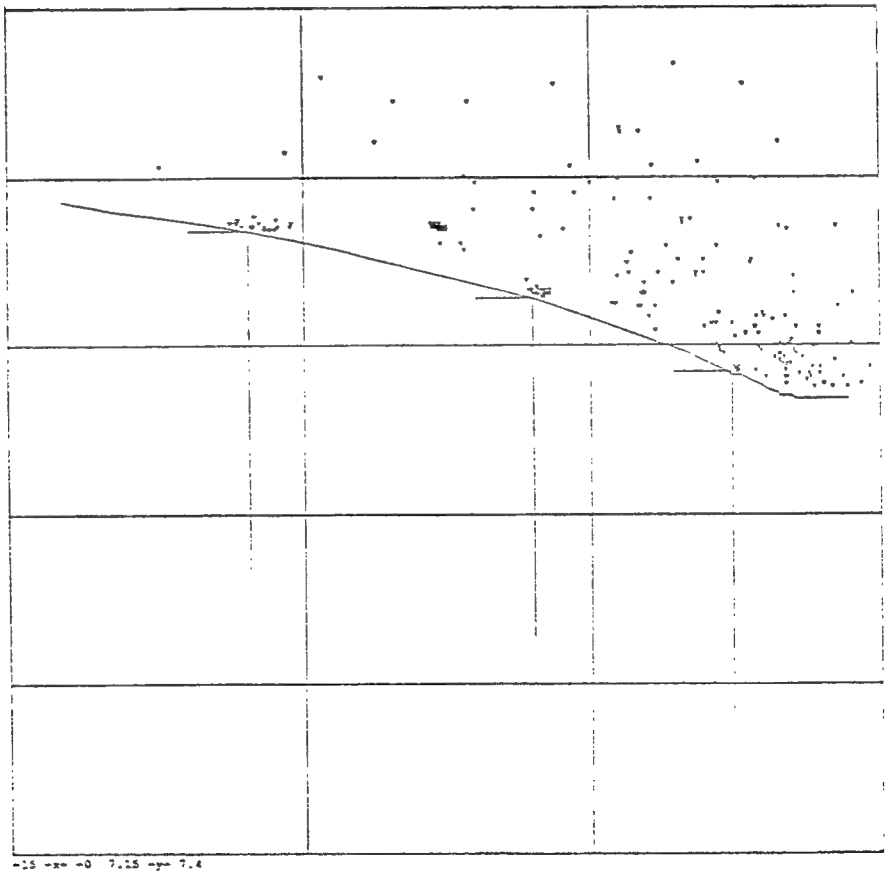


Fig. 2. Perturbations of minimal solutions

out numerically that this function $\varphi(Q_0)_{\tilde{\mu}}$ has several zeros that correspond to local Pareto minima. Table 1 contains the parameter $\tilde{\mu}$ with the corresponding $Q_0$ and SNR.

Figure 1 shows the different branches of solutions of these necessary conditions for this range of $\tilde{\mu}$-values. Numerical tests show that the branch of solutions corresponding to the lowest $Q$-value is likely to consist of Pareto minima. In Fig. 2 we show the distribution of $(-\text{SNR}, Q)$-values corresponding to complex perturbations of the individual components of the surface currents x associated with each of three $(-\text{SNR}, Q)$ points on the lowest branch. In each case, all the resulting $(-\text{SNR}, Q)$ pairs lie outside the negative cone as indicated in the figure. Hence, within the range of the perturbations, the points on the lowest branch appear to represent Pareto minima.

We note, finally, that this lower branch shows a relatively wide variation in the value of SNR for very small changes in the value of the quality factor $Q$. This indicates that one should be able to achieve relatively high values of SNR without an appreciable degradation of the quality factor.

## Acknowledgements

This work was supported under AFOSR Grant 91-0277 and DFG Grant Kr 940/1-1.

## References

1. Angell, T. S., Kirsch, A. and Kleinman, R. E., 'Antenna control and optimization', *Proc. IEEE*, 79, 1559–1568 (1991).
2. Colton, D. L. and Kress, R., *Integral Equation Methods in Scattering Theory*, Wiley, New York, 1983.
3. Dauer, J. P. and Stadler, W., 'A survey of vector optimization in infinite dimensional spaces, part II', *J. Optim. Theory and Appl.*, 51, 205–241 (1986).
4. Jahn, J., *Mathematical Vector Optimization in Partially Ordered Linear Spaces*, Peter Lang, Frankfurt, 1986.
5. Kirsch, A., Warth, W. and Werner, J., 'Notwendige Optimalitätsbedingungen und ihre Anwendung', *Lecture Notes in Economics and Mathematical Systems*, Vol. 152, Springer, Berlin, 1978.
6. Kirsch, A. and Wilde, P., 'The optimization of directivity and signal-to-noise ratio of an arbitrary antenna array', *Math. Meth. in the Appl. Sci.*, 10, 153–164 (1988).
7. Lo, Y. T., Lee, S. W. and Lee, Q. H., 'Optimization of directivity and signal-to-noise ratio of an arbitrary antenna array', *Proc. IEEE*, 54, 1033–1045 (1966).
8. Rhodes, D. R., *Synthesis of Planar Antenna Sources*, Clarendon, Oxford, 1974.
9. Stadler, W., 'A survey of multicriteria optimization or the vector maximization problem, part I: 1776–1960', *J. Optim. Theory and Appl.*, 29, 1–52 (1979).
10. Stadler, W., 'Multicriteria optimization in mechanics (a survey)', *Appl. Mech. Rev.*, 37, 277–286 (1984).
11. Woodward, P. M. and Lawson, J. D., 'The theoretical precision with which an arbitrary radiation pattern may be obtained from a source of finite size', *J. Inst. Elect. Engineers*, 95, 363–370 (1948).
12. Yosida, K., *Functional Analysis*, 4th edn, Springer, Berlin, 1974.

# THE REVIEW OF RADIO SCIENCE

## 1990–1992

*Edited by*

W. ROSS STONE

*Stoneware Limited*
*1446 Vista Claridad*
*La Jolla, California 92037-7839 USA*

*Associate Editor*

GEOFFREY HYDE

*Consultant*
*15116 Red Clover Drive*
*Rockville, MD 29853-1642 USA*

*URSI Commission Editors*

*Commission A:* Peter I. Somlo
*Commission B:* Staffan Ström
*Commission C:* Peter A. Matthews
*Commission D:* Tatsuo Itoh
*Commission E:* Pierre Degauque
*Commission F:* Richard K. Moore
*Commission G:* K. C. Yeh
*Commission H:* François Lefeuvre
*Commission J:* Tasso Tzioumis
*Commission K:* Paolo Bernardi

Published for the
International Union of Radio Science
by
Oxford University Press
1993

# 3. Iterative Methods for Radio-Wave Problems

*Ralph E. Kleinman*
Center for the Mathematics of Waves
Department of Mathematical Sciences
University of Delaware
Newark, DE 19716, USA

*Peter M. van den Berg*
Laboratory for Electromagnetic Research
Faculty of Electrical Engineering
Delft University of Technology
PO Box 5031, 2600 GA Delft, The Netherlands

## 1. Abstract

In recent years, there has been increasing interest in the use of iterative methods for solving a variety of problems in propagation, scattering, and inverse scattering of radio waves. In this paper, we review a number of the most-prominent methods for solving operator equations, arising in wave-field problems. The methods are used both in time-domain and frequency-domain problems and, in the frequency domain, are most useful at low and intermediate frequencies. In direct-scattering problems, we describe the essential features of the Neumann series, over-relaxation methods, Krylov-subspace methods, conjugate-gradient and biconjugate-gradient methods, and the conjugate-gradient-squared technique. Most of these methods are shown to be derivable from an error-minimization principle using various error criteria. Convergence of these methods is discussed. The error-minimization principle is shown to underlie a number of approaches to inverse problems, of reconstructing complex indices of refraction and scattering shape, from scattered-field measurements. The same iterative methods used in the (linear) direct problems are also applicable in the (nonlinear) inverse problems.

## 2. Introduction

Radio-wave problems are often formulated as integral equations, and it is this form which serves as the starting point for most numerical solutions. Typically, the integral operators which occur are boundary integrals, when considering scattering by impenetrable or penetrable homogeneous objects, and domain integrals, for penetrable inhomogeneous scatterers. These operators are invariably complex and non-self-adjoint, which complicates most numerical approaches. In a large number of

cases, the integral equations may be put in an abstract Hilbert-space framework, of the form

$$Lu = f, \tag{1}$$

where $L$ is a bounded, invertible linear operator, usually non-self-adjoint, which maps a Hilbert space, $H$, into another Hilbert space, $H'$. Best known are the cases when $L$ is a second-kind integral operator, of the form identity plus compact operator, but first-kind equations also fit into this framework. For example, *Van den Berg et al.* [*1991*] have shown that the first-kind operator, arising in scattering by a strip, is a boundedly-invertible map from the Sobolev space, $H_{-1/2}$ to $H_{1/2}$, and the operator equation may be solved iteratively. For simplicity, we will take $H = H'$; however, the iterative methods we describe may all be extended to this more general case. If $L^{-1}$ is not bounded, convergence of many of the iterative methods may still be established (e.g., *Browder and Petryshyn* [*1966*], *Brakhage* [*1987*]). However, convergence is much slower, and here we confine attention to the simpler case, where $L^{-1}$ is bounded. The space will be equipped with a norm, $\|\cdot\|$, and an inner product, $\langle\cdot,\cdot\rangle$, linear in the first entry. The adjoint operator is denoted as $L^*$. In this paper, we will consider iterative solutions of Equation (1) of the form

$$u_0 \text{ arbitrary}, \quad u_n = u_{n-1} + \alpha_n v_n, \ n \geq 1, \tag{2}$$

with the residual

$$r_n = f - Lu_n, \ n \geq 0 \tag{3}$$

Different choices of the constant $\alpha_n$, and of the update function or direction, $v_n$, give rise to different iterative schemes.

The literature on iterative methods is vast, and no attempt will be made to present a complete bibliography. A sampling of key references will be given in the sections describing particular schemes. Two excellent and up-to-date reviews of iterative methods for linear systems, *Ashby et al.* [*1990*] and *Freund et al.* [*1992*], deserve particular mention. Of special interest for the application of iterative methods in electromagnetics is the recent book, *Sarkar* [*1991*], Chapter three of which contains a more-detailed description of much of the material summarized here.

## 3. Neumann iteration

The simplest iterative procedure we consider results from the choice,

$$\alpha_n = 1, \quad v_n = r_{n-1} \tag{4}$$

This gives rise to the well known Neumann or Picard-Poincaré-Neumann iteration:



$$u_n = u_{n-1} + r_{n-1} = f + (I - L)u_{n-1}. \tag{5}$$

It is well known that a condition sufficient to ensure convergence of this iterative process is that the spectral radius of $I - L$ be less than one. Recently, a convergence criterion for the time-domain iterative Born approximation, in scattering by an inhomogeneous object, has been given by *De Hoop* [*1991*]

Most of the methods considered subsequently in this paper may be considered as generalizations of the Neumann iteration, based on more-elaborate choices of $\alpha_n$ and $v_n$. These choices will be guided by the desire to minimize either $\|u - u_n\|$, $\|L(u - u_n)\| = \|r_n\|$, or $\langle u - u_n, L(u - u_n)\rangle$ (the latter case only if $L$ is not only self-adjoint but also positive), on some subspace, $M$ of $H$. We define the functions $v_n$ to be

$$v_1 = Tr_0,$$

$$v_n = Tr_{n-1} + \sum_{m=1}^{n-1} \gamma_{n,m} v_m, \quad n > 1, \tag{6}$$

where $T$ is a linear operator, and $\gamma_{n,m}$ are constants, all yet to be chosen. Because of the iterative way in which the $v_n$ are constructed, it is easily seen that, for each $n$, $v_n$ is a linear combination of $\{T(LT)^{m-1} r_0;\ m = 1,\ldots,n\}$. Subspaces spanned by successive iterates of a single element, in this case $r_0$ (or $f$), are called Krylov subspaces generated by the element $r_0$ and the operator $LT$. Techniques for approximating solutions of operator equations in such subspaces are called Krylov-subspace methods.

## 4. Stationary over-relaxation method

The simplest generalization of the Neumann series is the stationary over-relaxation method, which results from the choices $\alpha_n = \alpha = \text{const}$ in Equation (2) and $\gamma_{n,m} = 0$ in Equation (6). Then, Equation (2) becomes

$$u_n = u_{n-1} + \alpha T r_{n-1} = \alpha Tf + (I - \alpha TL)u_{n-1}. \tag{7}$$

Convergence of this method is assured if the spectral radius of $I - \alpha LT$ is less than one. If $L$ is positive (and hence self-adjoint), then the spectrum of $L$ is real, and if Equation (1) is always uniquely solvable, then there always exists an $\alpha (0 < \alpha < \frac{2}{\|L\|})$ for which the iteration converges for $T = I$. If $L$ is not self-adjoint, then the existence of $\alpha$ for which the iteration scheme, with $T = L^*$, converges, is assured because of

unique solvability, and we have convergence if $0 < \alpha < \frac{2}{\|L^*L\|}$. In effect, we solve the normal equation $L^*Lu = L^*f$, rather than Equation (1), but they are equivalent if Equation (1) is uniquely solvable for all $f \in H$. While it is easier to establish convergence of this iterative scheme than when $T = I$, since less information is needed about the spectrum, the disadvantage is the appearance of the operator $L^*L$, which negatively affects the rate of convergence.

If $L$ is positive, an optimal $\alpha$ was found [*Kantorovich and Akilov, 1982*] to be

$$\alpha = \frac{2}{\|L\| + \|L^{-1}\|}, \tag{8}$$

which, of course, is useful only if good estimates of $\|L\|$ and $\|L^{-1}\|$ are available. The determination of an optimal $\alpha$ is quite difficult, in general. However, easily-computable and effective choices result from minimizing functionals of the first iterate. If $L$ is positive, then an easily-computable value may be obtained by minimizing $\langle u - u_1, L(u - u_1)\rangle$, with $T = I$, resulting in

$$\alpha = \frac{\|r_0^2\|}{\langle r_0, Lr_0\rangle}. \tag{9}$$

If $L$ is not self-adjoint then minimizing $\|u - u_1\|$, with $T = L^*$, leads to

$$\alpha = \frac{\|r_0^2\|}{\|L^* r_0\|^2}, \tag{10}$$

while minimizing $\|r_1\|$ with $T = I$ leads to

$$\alpha = \frac{\langle r_0, Lr_0\rangle}{\|Lr_0\|^2}. \tag{11}$$

The latter is preferable, since explicit occurrence of the adjoint, $L^*$, is avoided in the iterative scheme defined by Equation (7).

The method described in this section derives its name from the work of *Petryshyn* [*1962, 1963*], on generalized over-relaxation methods for operator equations, of which the present method is the simplest example. It is the operator equivalent of a one-step, or stationary-Richardson, method, in matrix theory. It was used by *Landweber* [*1954*] and *Fridman* [*1956*], when $L$ was a first-kind self-adjoint integral operator, and by *Bialy* [*1959*], for self-adjoint first- and second-kind operators. Applicability to the integral equations arising in scattering problems was



shown by *Kleinman and Roach* [*1988*]. Petryshyn's work also treated non-self-adjoint operators, and his general convergence criteria was found also by *Chertock* [*1968*], in the particular case of an integral equation in acoustic scattering. Convergence criteria and numerical examples are given by *Kleinman et al.* [*1990a, 1990b*].

## 5. Successive over-relaxation method

The next specialization of the general iterative method is the successive over-relaxation method, which results from again choosing $\gamma_{n,m} = 0$ in Equation (6), but letting $\alpha_n$ vary with $n$. Then, Equation (2) becomes

$$u_n = u_{n-1} + \alpha_n T r_{n-1}, \quad \text{and} \quad r_n = (I - \alpha_n LT) r_{n-1}. \tag{12}$$

We choose $\alpha_n$ so that $u_n$ minimizes some functional of $u - u_n$ on the subspace $M_n$, consisting of all linear combinations of $\{Tr_{m-1};\ m = 1,\ldots,n\}$.

If $L$ is positive, then minimizing $\langle u - u_n, L(u - u_n)\rangle$ with $T = I$ leads to

$$\alpha_n = \frac{\|r_{n-1}\|^2}{\langle r_{n-1}, Lr_{n-1}\rangle}. \tag{13}$$

If $L$ is not self-adjoint, then minimizing $\|u - u_n\|$ with $T = L^*$ leads to

$$\alpha_n = \frac{\|r_{n-1}\|^2}{\|L^* r_{n-1}\|^2}. \tag{14}$$

Alternatively, minimizing $\|r_n\|$ with $T = I$ yields

$$\alpha_n = \frac{\langle r_{n-1}, Lr_{n-1}\rangle}{\|Lr_{n-1}\|^2}, \tag{15}$$

whereas minimizing $\|r_n\|$ with $T = L^*$ yields

$$\alpha_n = \frac{\|L^* r_{n-1}\|^2}{\|LL^* r_{n-1}\|^2}. \tag{16}$$

An important difference between the successive and the stationary over-relaxation schemes is that while the existence of $\alpha$ such that $\|I - \alpha L\| < 1$ will guarantee convergence of the successive iteration, explicit determination of this value of $\alpha$ is

not needed in order to actually carry out the iterative procedure. This is in contrast with the stationary scheme where $\alpha$ is needed to implement the process.

All of the processes are convergent, provided $L$ is bounded and boundedly invertible, except for the case when $L$ is not self-adjoint, $\|r_n\|$ is minimized with $T = I$, and $\alpha_n$ is given by Equation (15). Then, additional conditions on $L$ are required to establish convergence: For example, the existence of $\alpha$, such that $\|I - \alpha L\| < 1$, will ensure convergence in this case. However, numerical experiments indicate that, while establishing convergence is most difficult in this case, this algorithm outperforms the other successive over-relaxation methods given here.

The method described here as successive over-relaxation is a natural generalization of the method described in the previous section. However, the matrix analog of the present method is called Richardson iteration, and the term "successive over-relaxation" in matrix theory involves a relaxation parameter, which does not change with each iteration, and more-nearly conforms with what is termed here as stationary over-relaxation. Numerical implementation of the present methods has been carried out by *Kleinman and Van den Berg* [see *Sarkar, 1990*, Chapter 3], for $L$ non-self-adjoint, for the cases described in Equation (14) - (16). The most rapid convergence was exhibited by the method associated with Equation (15), and this has been employed in electromagnetic scattering by *Samokhin* [*1990*]. The method associated with Equation (16) has been employed in a time-domain problem by *Herman and Van den Berg* [*1982*].

## 6. Krylov-subspace methods

The next method considered is the full Krylov method, by which we mean the general iterative method of Equation (2), with no prior restrictions on $\gamma_{n,m}$ in Equation (6) or on $\alpha_n$ in Equation (2). Thus, we consider the scheme

$$u_0 \quad \text{arbitrary},$$

$$u_n = u_{n-1} + \alpha_n v_n, \quad n > 1, \tag{17}$$

$$v_1 = Tr_0, \tag{18}$$

$$v_n = Tr_{n-1} + \sum_{m=1}^{n-1} \gamma_{n,m} v_m, \quad n > 1 \tag{19}$$

If $L$ is positive, then minimizing $\langle u - u_n, L(u - u_n) \rangle$ with $T = I$ leads to

$$\alpha_n = \frac{\|r_{n-1}\|^2}{\langle v_n, Lv_n \rangle} \text{ and } \gamma_{n,m} = -\frac{\langle r_{n-1}, Lv_m \rangle}{\langle v_m, Lv_m \rangle}. \tag{20}$$



If $L$ is not self-adjoint, then minimizing $\|u - u_n\|$ with $T = L^*$ leads

$$\alpha_n = \frac{\|r_{n-1}\|^2}{\|v_n\|^2} \text{ and } \gamma_{n,m} = -\frac{\langle L^* r_{n-1}, v_m \rangle}{\|v_m\|^2}. \tag{21}$$

Alternatively, minimizing $\|r_n\|$ with $T = I$ yields

$$\alpha_n = \frac{\langle r_{n-1}, Lr_{n-1} \rangle}{\|Lv_n\|^2} \text{ and } \gamma_{n,m} = -\frac{\langle Lr_{n-1}, Lv_m \rangle}{\|Lv_m\|^2}, \tag{22}$$

which is known as the generalized conjugate-residual (GCR) method, whereas minimizing $\|r_n\|$ with $T = L^*$ yields

$$\alpha_n = \frac{\|L^* r_{n-1}\|^2}{\|Lv_n\|^2} \text{ and } \gamma_{n,m} = -\frac{\langle LL^* r_{n-1}, Lv_m \rangle}{\|Lv_m\|^2}. \tag{23}$$

In the special case when $\langle Tr_n, Tr_m \rangle = 0$, $n \neq m$, the full Krylov method reduces to the successive over-relaxation method, and thus it will have the same convergence properties. Thus, all of the Krylov algorithms converge, provided $L$ is bounded and boundedly invertible, except for the GCR method. In that case, additional conditions on $L$ are needed, as described in the previous section. However, in general $\langle Tr_n, Tr_m \rangle \neq 0$, in which case it is expected that the full Krylov method will converge faster than the over-relaxation method. Of the various Krylov algorithms described, numerical evidence gives preference to the GCR method provided, of course, that it converges.

A number of sophisticated variants of Krylov-subspace methods have been derived in the numerical analysis community, for solving linear systems with non-symmetric coefficient matrices. These include ORTHODIR and ORTHORES [*Jea and Young, 1980*], ORTHOMIN [*Vinsome, 1976*] and, perhaps the best known, GMRES [*Saad and Schultz, 1986*]. Discussion of these methods and an extensive bibliography is found in the excellent review, *Freund et al.* [*1992*]. Although these methods have been applied in some electromagnetics problems, the most popular iterative method in electromagnetics, in the recent past, is the conjugate-gradient method, which we describe next.

## 7. Conjugate-gradient method

Considerable simplification occurs if $TL$ is self-adjoint, in which case we find that $v_n$ depends only on $r_{n-1}$ and $v_{n-1}$. Then, the Krylov-subspace method becomes the conjugate-gradient method. In particular, the algorithm becomes

$u_0$ arbitrary,

$$u_n = u_{n-1} + \alpha_n v_n, \quad n \geq 1, \tag{24}$$

$$r_n = f - Lu_n, \quad n \geq 0, \tag{25}$$

$$v_1 = Tr_0, \tag{26}$$

$$v_n = Tr_{n-1} + \gamma_{n,n-1} v_{n-1}, \quad n > 1. \tag{27}$$

The constants $\alpha_n$ and $\gamma_{n,n-1}$ are found by minimizing some functional of $u - u_n$ at each step.

If $L$ is positive, then by minimizing $\langle u - u_n, L(u - u_n)\rangle$ with $T = I$, we find

$$\alpha_n = \frac{\|r_{n-1}\|^2}{\langle v_n, Lv_n\rangle} \text{ and } \gamma_{n,n-1} = -\frac{\langle r_{n-1}, Lv_{n-1}\rangle}{\langle v_{n-1}, Lv_{n-1}\rangle} = \frac{\|r_{n-1}\|^2}{\|r_{n-2}\|^2}. \tag{28}$$

This is the original conjugate-gradient algorithm of *Hestenes and Stiefel [1952]*.

If L is self-adjoint and not necessarily positive, then by minimizing $\|r_n\|$ with $T = I$ we obtain

$$\alpha_n = \frac{\langle r_{n-1}, Lr_{n-1}\rangle}{\|Lv_n\|^2} \text{ and } \gamma_{n,n-1} = -\frac{\langle Lr_{n-1}, Lv_{n-1}\rangle}{\|Lv_{n-1}\|^2} = \frac{\langle r_{n-1}, Lr_{n-1}\rangle}{\langle r_{n-2}, Lr_{n-2}\rangle}. \tag{29}$$

If L is not self-adjoint, then minimizing $\|u - u_n\|$ with $T = L^*$ leads to

$$\alpha_n = \frac{\|r_{n-1}\|^2}{\|v_n\|^2} \text{ and } \gamma_{n,n-1} = -\frac{\langle L^* r_{n-1}, v_{n-1}\rangle}{\|v_{n-1}\|^2} = \frac{\|r_{n-1}\|^2}{\|r_{n-2}\|^2}, \tag{30}$$

whereas minimizing $\|r_n\|$ with $T = L^*$ yields

$$\alpha_n = \frac{\|L^* r_{n-1}\|^2}{\|Lv_n\|^2} \text{ and } \gamma_{n,n-1} = -\frac{\langle LL^* r_{n-1}, Lv_{n-1}\rangle}{\|Lv_{n-1}\|^2} = \frac{\|L^* r_{n-1}\|^2}{\|L^* r_{n-2}\|^2}, \tag{31}$$

which is the Hestenes-Stiefel conjugate-gradient algorithm applied to the normal equation, $L^* Lu = L^* f$.



A simplified form of the conjugate-gradient algorithms results from renormalizing the correction direction, $v_n$. If $\langle u - u_n, L(u - u_n)\rangle$ is minimized ($L$ self-adjoint and positive) and with $T = I$, then $v_n = \|r_{n-1}\|^2 w_n$, and the algorithm becomes

$$u_n = u_{n-1} + \frac{w_n}{\langle w_n, Lw_n\rangle}, \quad w_n = w_{n-1} + \frac{r_{n-1}}{\|r_{n-1}\|^2}. \tag{32}$$

If $\|r_n\|$ is minimized ($L$ self-adjoint, not necessarily positive) and $T = I$, then $v_n = \langle Lr_{n-1}, r_{n-1}\rangle w_n$, and the algorithm becomes

$$u_n = u_{n-1} + \frac{w_n}{\|Lw_n\|^2}, \quad w_n = w_{n-1} + \frac{r_{n-1}}{\langle r_{n-1}, Lr_{n-1}\rangle}. \tag{33}$$

If L is not self-adjoint and $T = L^*$, and if $\|u - u_n\|$ is minimized, then $v_n = \|r_{n-1}\|^2 w_n$, and the algorithm becomes

$$u_n = u_{n-1} + \frac{w_n}{\|w_n\|^2}, \quad w_n = w_{n-1} + \frac{L^* r_{n-1}}{\|L^* r_{n-1}\|^2}, \tag{34}$$

which is the method of Le Foll [*Le Foll, 1971*; *Lardy, 1990*]. Whereas, if $\|r_n\|$ is minimized, then $v_n = \|L^* r_{n-1}\|^2 w_n$, and the algorithm becomes

$$u_n = u_{n-1} + \frac{w_n}{\|Lw_n\|^2}, \quad w_n = w_{n-1} + \frac{L^* r_{n-1}}{\|L^* r_{n-1}\|^2}. \tag{35}$$

When $TL$ is self-adjoint and positive, the conjugate-gradient method will always converge. Thus, all of the algorithms given here converge, except for that case corresponding to the choice given in either Equation (29) or Equation (33). Then, additional restrictions on $L$ are needed, as previously described.

The literature on the conjugate-gradient method is vast, and no attempt is made to summarize it. An excellent review on work up to 1976 is given by *Golub and O'Leary [1989]*. More recent work is summarized by *Ashby et al. [1990]*. Applications in electromagnetics are reviewed in *Sarkar [1991]*, e.g., Woodworth and Yaghjian (Chapter 4), Wilton and Wheeler (Chapter 5), and Tijhuis (Chapter 13).

One of most popular methods in electromagnetics is the conjugate-gradient fast Fourier transform (CGFFT), which has been applied to various electromagnetic-field problems, where the conjugate-gradient method is combined with a fast Fourier transform technique for an efficient computation of the operator expressions with convolution kernels. Extensive explanation of the various forms, with many examples, can be found in *Sarkar [1991]*, viz., Van den Berg (Chapter 2), Volakis

and Barkeshli (Chapter 6), Peterson et al. (Chapter 7), Shen (Chapter 9), Lesselier and Duchêne (Chapter 10), and Ponnapali (Chapter 11). It should be mentioned that a direct replacement of the spatial derivatives in the integral operator, by simple algebraic multiplications in the spectral Fourier domain, may lead to substantial numerical difficulties and erroneous results, as shown by *Zwamborn and Van den Berg [1991a]*. Therefore, Gago and Cátedra [see *Sarkar, 1991*, Chapter 8] have introduced a careful testing and weighting procedure in the discretization of the integral equation, and solved the resulting equation by the CGFFT method. A more efficient procedure is the weak form of the CGFFT method, introduced by *Zwamborn and Van den Berg [1991a, 1991b]*. This method is denoted as the WCG-FFT method by *Bokhari et al. [1992]*, who use it for the computation of the radiation of microstrip antennas.

## 8. Biconjugate-gradient method

Considerable effort has been spent on alternative iterative methods for non-self-adjoint operators, which have less severe storage requirements than the full Krylov-subspace method and, at the same time, do not square the condition number, as in the conjugate-gradient method (with $T = L^*$). One such method is the biconjugate-gradient method [*Fletcher, 1976*], which involves augmenting Equation (1) with a second equation,

$$L^* \tilde{u} = \tilde{f}, \quad \text{with } \tilde{f} \text{ arbitrary,} \tag{36}$$

and considering the system of Equations (1) and (26), which may be written as

$$\mathcal{L}\vec{u} = \vec{f}, \quad \text{with } \mathcal{L} = \begin{pmatrix} 0 & L \\ L^* & 0 \end{pmatrix}, \ \vec{u} = \begin{pmatrix} \tilde{u} \\ u \end{pmatrix}, \ \vec{f} = \begin{pmatrix} f \\ \tilde{f} \end{pmatrix} \tag{37}$$

The biconjugate-gradient algorithm was originally proposed for the case where $L$ was a matrix, but applies equally well if $L$ is an integral operator. The operator $\mathcal{L}$ is self-adjoint, in the sense that $\langle\langle \vec{v}, \mathcal{L}\vec{u} \rangle\rangle = \langle\langle \mathcal{L}\vec{v}, \vec{u} \rangle\rangle$, where $\langle\langle \vec{v}, \vec{u} \rangle\rangle = \langle \tilde{v}, \tilde{u} \rangle + \langle v, u \rangle$ but, unfortunately, $\mathcal{L}$ is not positive definite. Using the iteration

$$\vec{u}_0 \quad \text{arbitrary,}$$

$$\vec{u}_n = \vec{u}_{n-1} + \alpha_n \vec{v}_n, \quad n \geq 1, \tag{38}$$

$$\vec{r}_n = \vec{f} - \mathcal{L}\vec{u}_n, \quad n \geq 0, \tag{39}$$

$$\vec{v}_1 = \mathcal{T}\vec{r}_0, \tag{40}$$



$$\vec{v}_n = \mathcal{T}\vec{r}_{n-1} + \sum_{m=1}^{n-2} \gamma_{n,m} \vec{v}_m, \quad n > 1, \tag{41}$$

all of the previously-described iteration schemes may be reproduced in vector form, yielding a family of biconjugate-gradient methods. However, what is generally known as the biconjugate-gradient algorithm, and the only one which exhibits rapid convergence in numerical examples, results from determining the constants $\alpha_n$ and $\gamma_{n,m}$ from formally minimizing $\langle\langle \vec{u} - \vec{u}_n, \mathcal{L}(\vec{u} - \vec{u}_n) \rangle\rangle$, with $\mathcal{T} = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix}$, which yields

$$\alpha_n = \frac{\langle\langle \vec{r}_{n-1}, \mathcal{T}\vec{r}_{n-1} \rangle\rangle}{\langle\langle \vec{v}_n, \mathcal{L}\vec{v}_n \rangle\rangle}, \ \gamma_{n,m} = 0, \ m < n-1, \ \gamma_{n,n-1} = \frac{\langle\langle \vec{r}_{n-1}, \mathcal{T}\vec{r}_{n-1} \rangle\rangle}{\langle\langle \vec{r}_{n-2}, \mathcal{T}\vec{r}_{n-2} \rangle\rangle} \tag{42}$$

The term $\langle\langle \bullet, \mathcal{L} \bullet \rangle\rangle$ is not a norm, since $\mathcal{L}$ is not positive definite, which places the procedure under a cloud, since convergence is not guaranteed. Nonetheless, the algorithm is remarkably effective in a large number of cases, although there are examples when it does not converge. No a priori test of convergence is available. A generalization of the biconjugate-gradient method, which avoids the breakdowns of that method, called the quasi-minimal-residual (QMR) algorithm, has recently been proposed [*Freund and Nachtigal, 1992*], but is yet to be tested on electromagnetic problems.

## 9. Conjugate-gradient-squared method

A variant of the biconjugate-gradient method which, in a sense, combines two iterative steps with the advantage of eliminating the operations involving $L^*$, has been proposed [*Sonneveld, 1989*]. Ingenious manipulation results in the following algorithm which, to date, has not been shown to be derivable by minimization of a functional of $u - u_n$:

$$u_0 \quad \text{arbitrary,} \ \tilde{r}_0 \neq 0, \ y_0 = 0, \ v_0 = 0,$$

$$u_n = u_{n-1} + \alpha_n (x_n + y_n), \quad n \geq 1, \tag{43}$$

$$r_n = f - L u_n, \quad n \geq 0, \tag{44}$$

$$x_n = r_{n-1} + \beta_n y_{n-1}, \quad n \geq 1, \tag{45}$$

$$y_n = x_n - \alpha_n L v_n, \quad n \geq 1, \tag{46}$$

$$v_n = x_n + \beta_n (y_{n-1} + \beta_n v_{n-1}), \quad n \geq 1, \tag{47}$$

where $\tilde{r}_0$ is suitably chosen, and

$$\alpha_n = \frac{\mathrm{Re}\langle \tilde{r}_0, r_{n-1}\rangle}{\mathrm{Re}\langle \tilde{r}_0, Lv_n\rangle}, \quad \beta_1 = \mathrm{Re}\langle \tilde{r}_0, r_0\rangle, \quad \beta_n = \frac{\mathrm{Re}\langle \tilde{r}_0, r_{n-1}\rangle}{\mathrm{Re}\langle \tilde{r}_0, r_{n-2}\rangle}, \quad n > 1. \tag{48}$$

Known as the conjugate-gradient-squared method, it has also been shown to be numerically effective in some cases. However, useful sufficient conditions on $L$ to guarantee convergence have not been found, and examples exist where the method fails to converge, even when the biconjugate-gradient method converges. When it does converge, however, it outperforms all other methods described. Some very recent work has appeared on avoiding breakdowns in the conjugate-gradient-squared method [*Brezinski and Sadok, 1991*] as well as on a variant of conjugate-gradient squared, based on quasi-error minimization [*Freund, 1992b*]. Yet another variant of the CGS (and BI-CG) method, which removes some of the irregular convergence behavior, has been proposed by *Van der Vorst* [*1992*]. Applications to electromagnetic problems have yet to be carried out, but steps in this direction have been taken [*Freund, 1992a*].

## 10. Multigrid methods

No review of iterative methods would be complete without some mention of multigrid techniques. These have received considerable attention in the past ten years, in fluid mechanics, but have yet to achieve widespread application in electromagnetics. Basic principles may be found in work by *Brandt* [*1977*] and *Hackbush* [*1985*]. Particularly relevant to the present review is the conclusion of *Wesseling* [*1992*] that for medium-sized linear problems, conjugate-gradient and multigrid methods are equally efficient, but conjugate-gradient methods are much easier to program.

## 11. Preconditioners

Preconditioners are usually thought of as approximate inverses, in the sense that a preconditioned operator is a perturbation of the identity. All iterative methods benefit from effective preconditioning, by reducing the amount of computation needed to achieve a preset error criterion. Preconditioners may be introduced either directly, by considering the equation $PLu = Pf$, where $P$ is the preconditioner, or by incorporating the preconditioner into the operator $T$ in the algorithms described above. *Van den Berg et al.* [*1991*] present a preconditioner for first-kind integral equations, arising in scattering by planar structures, which has proven remarkably effective in accelerating convergence of the CGFFT method (see also chapter 2 in *Sarkar* [*1991*]).

After discretization and reduction to a linear system, a variety of preconditioning techniques exists, usually involving the decomposition of the



coefficient matrix, e.g. *Axelsson* [*1985*]. Such methods are routinely incorporated into software packages implementing various conjugate-gradient algorithms: GCR, GMRES, ORTHORES, etc.

Methods that reduce the work in computing the coefficient matrix, or transform a dense matrix into a sparse one, may also be considered to be forms of preconditioning. In this regard, mention should be made of the fast-multipole technique (e.g. *Rokhlin* [*1990*]), and of the impedance-matrix-localization method [*Canning, 1990*].

## 12. Inverse problems

Iterative methods are also employed in most attacks on inverse problems: see, e.g., *Chew and Wang* [*1990*], Roger and Chapel [*Sarkar 1991*, Chapter 12]. One class of such problems consists of attempting to find an index of refraction profile, $\chi$, of an inhomogeneous object, $D$, which is successively irradiated by a number ($i = 1,\ldots,I$) of known fields, from the equations

$$u_i - G\chi u_i = f_i, \text{ in } D, \qquad Ku_i\chi = h_i, \text{ in } S, \tag{49}$$

where $G$ and $K$ are integral operators, $f_i$ is given in $D$, $h_i$ is measured on a surface, $S$, enclosing $D$, and $u_i$ as well as $\chi$ are unknown. In this profile-inversion problem, $u_i$ and $\chi$ are sought simultaneously, to minimize the $L^2$ error on $D$ and the $L^2$ error on $S$.

*Kleinman and Van den Berg* [*1992a*] proposed an iterative-inversion algorithm, which incorporates the ideas of successive over-relaxation as well as the conjugate-gradient method. Specifically, they propose the construction of sequences, $\{u_{i,n}\}$ and $\{\chi_n\}$ as follows(compare Equations (2) and (3)):

$$\begin{aligned} u_{i,0} &= u_i^{initial}, & \chi_0 &= \chi^{initial}, \\ u_{i,n} &= u_{i,n-1} + \alpha_n v_{i,n}, & \chi_n &= \chi_{n-1} + \beta_n d_n, \\ r_{i,n} &= f_i - G\chi_n u_{i,n}, & \rho_{i,n} &= h_i - Ku_{i,n}\chi_n, \end{aligned} \tag{50}$$

where $\alpha_n$ and $\beta_n$ are, in general, complex constants, which are chosen at each step to minimize

$$F_n = w_D \sum_{i=1}^{I} \|r_{i,n}\|_D^2 + w_S \sum_{i=1}^{I} \|\rho_{i,n}\|_S^2,$$

$$w_D = \left( \sum_{i=1}^{I} \left\| u_i^{inc} \right\|_D^2 \right)^{-1}, \quad w_S = \left( \sum_{i=1}^{I} \| f_i \|_S^2 \right)^{-1}, \tag{51}$$

and the subscripts $D$ and $S$, on the norm, $\|\cdot\|$, and inner product, $\langle \cdot, \cdot \rangle$, in $L^2$, indicate the domain of integration. The minimization of the quantity $E_n'$ of Equation (51) leads to a nonlinear problem for the coefficients $\alpha_n$ and $\beta_n$ at each step, which is solved using a conjugate-gradient method. The update direction for the field, $u_{i,n}$, is directly adapted from the successive over-relaxation method for solving the direct problem, and the update direction for the contrast, $\chi_n$, is taken to be the gradient of the error in the measured data. This inversion algorithm provided accurate profile reconstructions for low contrasts ($\chi \sim 1$). The range of contrasts successfully reconstructed was considerably enlarged by using more-sophisticated choices for the updating directions, based on the Polak-Ribière conjugate gradients [*Kleinman and Van den Berg, 1992b*]

## 13. Conclusions

We have reviewed a number of iteration schemes, based on error minimization, which are gradient methods. They all are generalizations of the Neumann series, in which the error is minimized over some subspace. The methods include stationary over-relaxation, in which the relaxation parameter is found by minimizing the residual error in the first iteration step; successive over-relaxation, in which the error at each step is minimized; and Krylov-subspace methods, in which the error is minimized over a subspace of all previous errors. Different definitions of the error lead to a number of different schemes, including conjugate-gradient algorithms.

Numerical examples indicate that better results come from more-sophisticated procedures, with the best performance shown by the Krylov method. However, in electromagnetic applications, by far the most popular iterative technique is the conjugate-gradient method. It has the advantage of requiring only the two previous residuals (or update directions) at each step, convergence is assured, and CG routines are readily available in many mathematical software packages. It has the disadvantage of requiring the use of the adjoint operator, since most operators arising in electromagnetics are not self-adjoint, and the consequent squaring of the condition number. These disadvantages are avoided in the Krylov method, which explains why it behaves better in the sense that it requires a lower iteration count than the conjugate-gradient method, for given accuracy. However, the disadvantage of the Krylov method is the necessary storage of all the update directions generated in previous iterations. The implicit orthogonalization in the conjugate-gradient methods overcomes this problem. A surprising result is that if orthogonality, which leads to the conjugate-gradient method, is ignored, and the full Krylov scheme is used even when not theoretically necessary, the results obtained are markedly superior to those obtained by the conjugate-gradient method, when more than a few iterations are employed. This reinforces the widely-held belief that in order to be effective, the Krylov method and the conjugate-gradient algorithm must be used together with a good preconditioner, which keeps the number of iterations small. However, all the methods described benefit from effective preconditioning. In order to decrease the number of necessary iterations, biconjugate-gradient methods and conjugate-gradient-squared methods have been introduced. However, useful sufficient conditions to guarantee convergence of the latter methods have not been found.



We have also discussed the (nonlinear) inverse problem. It is pointed out that iterative techniques, similar to those used in the linear direct problem, may also be used advantageously in the inverse problem.

## 14. Acknowledgment

This work was supported under NSF Grant No. DMS-8912593, AFOSR Grant-91-0277, ONR Grant N00014-91-J-1700, and NATO Grant-0230/88, and a Research Grant from Schlumberger-Doll Research, Ridgefield, CT, USA

## 15. References

S. F. Ashby, T. A. Manteuffel, and P. E. Saylor [1990], "A taxonomy for conjugate gradient methods," *SIAM J. Numer. Anal.*, **27**, pp. 1542-1568.

O. Axelsson [1985], "A survey of preconditioned iterative methods for linear systems of algebraic equations," *BIT*, **25**, pp. 166-187.

H. Bialy [1959], "Iterative behandlung linearer funktiongleichungen," *Arch. Rational Mech. Anal.*, **4**, pp. 166-176.

S. A. Bokhari, J. R. Mosig, and F. E. Gardiol [1992], "Radiation pattern computation of microstrip antennas on finite size ground planes," *Proc. IEE H*, **139**, pp. 278-286.

H. Brakhage [1987], "On ill-posed problems and the method of conjugate gradients," in H. W. Engl and C. W. Groetsch (eds.), *Inverse and Ill-posed Problems*, Boston, Academic Press, pp. 165-175.

A. Brandt [1977], "Multi-level adaptive solutions to boundary value problems," *Math. Comput.*, **31**, pp. 333-390.

C. Brezinski and H. Sadok [1991], "Avoiding breakdown in the CGS algorithm," *Numer. Algorithms*, **1**, pp. 199-206.

F. E. Browder and W. V. Petryshyn [1966], "The solution by iteration of linear functional equations in Banach spaces," *Bull. Amer. Math. Soc.*, **72**, pp. 566-570.

F. Canning [1990], "The impedance matrix localization (IML) method for moment method calculations," *IEEE Antennas and Propagation Magazine*, **32**, pp. 18-30.

G. Chertock [1968], "Convergence of iterative solutions to integral equations for sound radiation," *Q. Appl. Math.*, **XXVI**, pp. 269-272.

W. C. Chew and Y. M. Wang [1990], "Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method," *IEEE Trans. Med. Imag.*, **MI-9**, pp. 218-225.

A. T. de Hoop [1991], "Convergence criterion for the time-domain iterative Born approximation to scattering by an inhomogeneous, dispersive object," *J. Opt. Soc. Am.*, **A8**, 1256-1260.

R. Fletcher [1976], "Conjugate gradient methods for indefinite systems," in G. A. Watson (ed.), *Numerical Analysis Dundee 1975* (Lecture Notes in Mathematics 506), Berlin, Springer, pp. 73-89.

R. W. Freund [1992a], "Conjugate gradient-type methods for linear systems with complex symmetric coefficient matrices," *SIAM J. Sci. Stat. Comput.*, **13**, pp. 425-448.

R. W. Freund [1992b], "A transpose-free quasi-minimal residual algorithm for non-Hermitean linear systems," *SIAM J. Sci. Stat. Comput.*, (to appear).

R. W. Freund, G. H. Golub, and N. M. Nachtigal [1992], "Iterative solution of linear systems," *Acta Numerica*, (to appear).

R. W. Freund and N. M. Nachtigal [1992], "QMR: a quasi-numerical residual method for non-Hermitean linear systems," *Numer. Math.* (to appear).

V. Fridman [1956], "Method of Successive Approximations for a Fredholm Integral Equation of the First Kind," *Uspehi Mat. Nauk.*, **11**, pp. 233-234.

G. H. Golub and D. P. O'Leary [1989], "Some history of the conjugate gradient and Lanczos algorithms 1948-1976," *SIAM Review*, **31**, pp. 50-102.

W. Hackbush [1985], *Multi-grid Methods and Applications*, Berlin, Springer.

G. C. Herman and P. M. van den Berg [1982], "A least-square iterative technique for solving time-domain scattering problems," *J. Acoust. Soc. Am.*, **72**, pp. 1947-1953.

M. R. Hestenes and E. Stiefel [1952], "Methods of conjugate gradients for solving linear systems," *J. Res. Nat. Bur. Stand.*, **49**, pp. 409-436.



K. C. Jea and D. M. Young [1980], "Generalized conjugate gradient acceleration of nonsymmetrizable iterative methods," *Linear Algebra Appl.*, **34**, pp. 159-194.

L. V. Kantorovich and G. P. Akilov [1982], *Functional Analysis, 2nd ed.*, Oxford, Pergamon Press, p. 466.

R. E. Kleinman and G. F. Roach [1988], "Iterative solutions of boundary integral equations in acoustics," *Proc. Roy. Soc. London*, **A417**, pp. 45-57.

R. E. Kleinman, G. F. Roach, L. S. Schuetz, J. Shirron, and P. M. van den Berg [1990a], "An over-relaxation method for the iterative solution of integral equations in scattering problems," *Wave Motion*, **12**, pp. 161-170.

R. E. Kleinman, G. F. Roach, and P. M. van den Berg [1990b], "A convergent Born series for large refractive indices," *J. Opt. Soc. Am.*, **A7**, pp. 890-897.

R. E. Kleinman and P. M. van den Berg [1992a], "A modified gradient method for two-dimensional problems in tomography," *J. Comp. Appl. Math.*, **42**, pp. 17-35.

R. E. Kleinman and P. M. van den Berg [1992b], "An extended range modified gradient technique for profile inversion," *Proceedings of the 1992 URSI International Symposium on Electromagnetic Theory*, Sydney, pp. 251-253.

L. Landweber [1954], "An iteration formula for Fredholm integral equations of the first kind," *Amer. J. Math.*, **73**, pp. 615-624.

L. J. Lardy [1990], "A class of iterative methods of conjugate gradient type," *Numer. Funct. Anal. and Optimiz.*, **11**, pp. 283-302.

J. Le Foll [1971], "An iterative procedure for the solution of linear and nonlinear equations," in *Conference on Applications of Numerical Analysis* (Lecture Notes in Mathematics), New York, Springer-Verlag.

W. V. Petryshyn [1962], "The generalized overrelaxation method for the approximate solution of operator equations in Hilbert space," *J. Soc. Indust. Appl. Math.*, **10**, pp. 675-690.

W. V. Petryshyn [1963], "On a general iterative method for the approximate solution of linear operator equations," *Math. Comp.*, **17**, pp. 1-10.

V. Rokhlin [1990], "Rapid solution of integral equations of scattering theory in two dimensions," *J. Comp. Phys.*, **86**, pp. 414-439.

Y. Saad and M. H. Schultz [1986], "GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems," *SIAM J. Sci. Stat. Comput.*, **7**, pp. 856-869.

A. B. Samokhin [1990], "Analysis of problems of electromagnetic wave diffraction in inhomogeneous media," *USSR Computat. Math. and Math. Physics*, **30**, pp. 107-121.

T. K. Sarkar (ed.) [1991], *Application of Conjugate Gradient Method to Electromagnetics and Signal Analysis, Progress in Electromagnetic Research, PIER 5*, New York, Elsevier.

P. Sonneveld [1989], "CGS, a fast Lanczos-type solver for nonsymmetric linear systems," *SIAM J. Sci. Stat. Comput.*, **10**, pp. 36-52.

P. M. van den Berg, A. P. M. Zwamborn, G. C. Hsiao, and R. E. Kleinman [1991], "Iterative solution of first kind integral equations," in R. E. Kleinman, R. Kress, E. Martensen (eds.), *Direct and Inverse Boundary Value Problems*, Frankfurt am Main, Peter Lang, pp. 213-232.

H. A. van der Vorst [1992], "BI-CGSTAB: A fast and smoothly converging variant of BI-CG for the solution of non-symmetric linear systems," *Siam J. Sci. Stat. Comput.*, **13**, pp. 631-644.

P. K. W. Vinsome [1976], "ORTHOMIN, An iterative method for solving sparse sets of simultaneous linear equations," in *Proceedings of the Fourth Symposium on Reservoir Simulation*, Society of Petroleum Engineers of the American Institute of Mechanical Engineers, pp. 149-159.

P. Wesseling [1992], *An Introduction to Multigrid Methods*, Chichester, John Wiley & Sons, pp. 206-207.

A. P. M. Zwamborn and P. M. van den Berg [1991a], "A weak form of the conjugate gradient FFT method for plate problems," *IEEE Trans. Ant. Prop.*, AP-39, pp. 224-228.

A. P. M. Zwamborn and P. M. van den Berg [1991b], "A weak formulation of the conjugate gradient FFT method for two-dimensional TE scattering problems," *IEEE Trans. Microw. Theory Tech.*, MTT-39, pp. 953-960.

# Electromagnetic Scattering by Indented Screens

John S. Asvestas, *Senior Member, IEEE*, and Ralph. E. Kleinman, *Senior Member, IEEE*

*Abstract*—**The problem of three dimensional electromagnetic scattering from a perfectly conducting screen with a bounded indentation is formulated as a system of boundary integral equations for the electric current density on the cavity wall and the interface between the cavity and free space. It is shown how the fictitious current density on the interface may be eliminated resulting in an integral equation of the second kind for the current density on the cavity wall only, with no integration over the infinite screen. In addition, integral representations are derived that represent the field everywhere in space in terms of the current density on the cavity wall only. Furthermore, asymptotic expressions for the far field are also presented. The equations and representations simplify considerably in the two–dimensional scalar case and results are presented for both TE and TM polarization.**

## I. Introduction

BOUNDARY integral equation formulations of electromagnetic scattering problems serve as one of the primary bases for numerical approximations. The electric and magnetic field integral equation formulations of scattering of an incident field from a bounded obstacle immersed in free space are well known (e.g., Poggio and Miller [1], Colton and Kress [2]) as are alternate forms which have been developed to eliminate illposedness (non-uniqueness) of these equations at interior eigenvalues or cavity resonances. These include the combined field and combined source equations (e.g., Brakhage and Werner [3], Burton and Miller [4], Harrington and Mautz [5]) and modified Green's function methods (e.g., Jones [6], Kleinman and Roach [7], Jost [8], and Yaghjian [9]).

When the scattering object is above a conducting plane, the integral equation formulation is much the same. Using the method of images, integrals over the plane may be removed by introducing the Green's function for the plane as the fundamental solution. The resulting integral equations are changed only by replacing the free space Green's function by the Green's function for the plane and adding a reflected field to the known term.

When the scattering object punctures the conducting plane, the problem is still easily reduced to familiar integral equations. In the case of scattering by extrusions on a perfectly conducting plane, the presence of the plane may be taken into account by combining the results for scattering from an unperturbed plane with the field scattered by an obstacle consisting of the extrusion and its image in the plane, a bounded obstacle in free space.

The picture is dramatically changed when the conducting plane has indentations rather than extrusions. The previously cited methods of reduction to simpler problems which do not involve integrals over the infinite plane are no longer available. Moreover this case has become important since it has been observed that small indentations may considerably change the scattering characteristics of otherwise smooth surfaces.

This problem has received considerable attention over the years. Most of the boundary integral formulations in the engineering literature are based on Schelkunoff's equivalence principle, which is essentially an application of the vector Green's theorem, see Chen [10], coupled with the network formulation of Mautz and Harrington [11]. The scattering domain is decomposed into two parts, an infinite half space and a cavity in the plane. The two are connected via currents on the fictitious surface between the cavity and the half space. However, the integral equation for the closed domain bounded by the actual physical indentation and the fictitious surface separating the cavity from the half space is plagued by the usual problem of non-uniqueness at the cavity resonances, e.g., Liang and Cheng [12]. In this approach the fields interior and exterior to the cavity are coupled by a fictitious magnetic current on the interface between the cavity and free space. The virtue of this equation is its relative simplicity; however, the price one pays is the occurrence of spurious resonances.

Recently a different attack has been made using a combined finite element boundary element approach wherein the boundary integral equation arising from the exterior half space is coupled on the fictitious surface with a finite element formulation for the fields in the cavity (see Jin and Volakis [13, 14], Jeng [15], and Jeng and Tzeng [16]). A variation on this approach was used by Wang and Ling [17, 18] in which the cavity was decomposed into subelements each of which was treated via integral equations on the subelement boundaries. These hybrid approaches apparently eliminate the resonance problem (although no theoretical uniqueness proof is available yet) at the cost of a finite element rather than boundary element computation.

For related scalar problems in acoustic scattering, Willers [19, 20] derived boundary integral equations for unknown functions defined only on the surface of the indentation and showed that these equations were uniquely solvable for all $k$ for both Dirichlet and Neumann boundary conditions. However the kernel of Willer's integral equations involve integrals of free-space Green's functions over the entire boundary, screen plus indentation, and therefore are awkward for numerical implementation.

Manuscript received June 15, 1992; revised June 29, 1993. This work was supported in part under AFOSR Grant 91–0277 and ONR Grant N00014–91–J–1700.

J. S. Asvestas is with the Grumman Corporate Research Center, M/S A01–26, Bethpage, NY 11714-3580.

R. E. Kleinman is with the Center for the Mathematics of Waves, Department of Mathematical Sciences, University of Delaware, Newark, DE 19716.

IEEE Log Number 9215046.

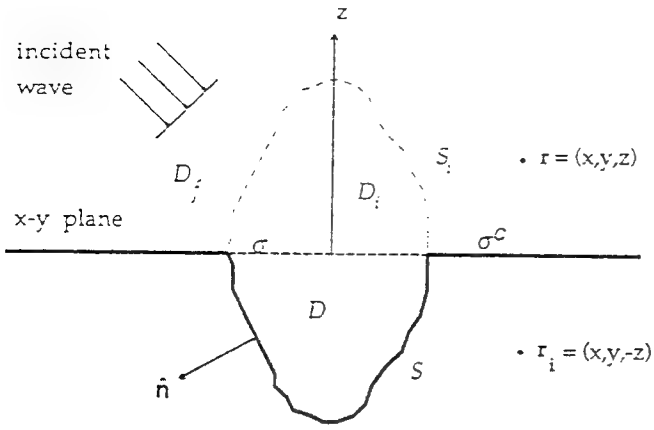



Fig. 1. Geometry of the Indented Screen Problem.

In the present paper we develop boundary integral equations for the electromagnetic indented screen problem that avoid the introduction of a "magnetic" current density. We obtain integral equations in terms of an electric current density on both the cavity wall and the interface between the cavity and free space, and show how these lead to a second kind integral equation for the electric current density on the cavity boundary only. This approach is based on the direct approach using Green's theorem rather than a layer ansatz. The integral equations obtained do not involve any integrals over the infinite screen and, thus, lend themselves to numerical computation. Moreover, these equations are expected to be free of cavity resonances though this is still to be proved.

In Section II we state the problem and in Section III the main results: an integral equation for the current density on the cavity wall and integral representations of the fields everywhere in space in terms of the solution of the integral equation. These representations simplify considerably in the far field and the results are presented in Section IV. Section V contains the corresponding two-dimensional integral equations and representations for both transverse electric and transverse magnetic polarizations. The derivation of the main results is presented in the Appendix.

## II. Statement of the Problem

The geometry of the problem is shown in Fig. 1. The domain of interest is that exterior to an indented perfectly conducting plane. The plane is taken to coincide with the $x$-$y$ plane in a Cartesian coordinate system. The bounded indentation, denoted by $D$, has a boundary consisting of two parts: $S$ which lies in the lower half space and $\sigma$, a portion of the $x$-$y$ plane. The entire plane consists of $\sigma$ and its unbounded complement $\sigma^c$. In the scattering problem the boundary consists of $S$ and $\sigma^c$.

We consider an incident electromagnetic field $(\mathbf{E}^{\text{inc}}(\mathbf{r}), \mathbf{H}^{\text{inc}}(\mathbf{r}))$ originating in the upper half space, with the restriction that no sources exist in the image of $D$ or its boundary and assume that all field quantities have a harmonic time dependence, $e^{-i\omega t}$, which is suppressed. The subscript $i$ on a vector quantity indicates its image in the $x$-$y$ plane, for example if the Cartesian components of $\mathbf{E}(\mathbf{r})$ are given by

$$\mathbf{E}(\mathbf{r}) = (E_x(\mathbf{r}), E_y(\mathbf{r}), E_z(\mathbf{r})), \tag{2.1}$$

then by $\mathbf{E}_i(\mathbf{r})$ we mean

$$\mathbf{E}_i(\mathbf{r}) = (E_x(\mathbf{r}), E_y(\mathbf{r}), -E_z(\mathbf{r})) \tag{2.2}$$

Consistent with this notation we denote a position vector by $\mathbf{r} = (x, y, z)$ and its image in the $x$-$y$ plane by $\mathbf{r}_i = (x, y, -z)$. Moreover we denote by $D_i$ and $S_i$ the images of $D$ and $S$ in $x$-$y$ plane and by $D_f$ that part of the upper half space excluding $D_i$ and $S_i$.

If the indentation is absent, so that we are treating scattering by a perfectly conducting plane, then the total field may be found by the method of images to be

$$\begin{aligned}\mathbf{E}^0(\mathbf{r}) &:= \mathbf{E}^{\text{inc}}(\mathbf{r}) - \mathbf{E}_i^{\text{inc}}(\mathbf{r}_i)\\ \mathbf{H}^0(\mathbf{r}) &:= \mathbf{H}^{\text{inc}}(\mathbf{r}) + \mathbf{H}_i^{\text{inc}}(\mathbf{r}_i)\end{aligned} \tag{2.3}$$

Since $(\mathbf{E}^{\text{inc}}(\mathbf{r}), \mathbf{H}^{\text{inc}}(\mathbf{r}))$ satisfy the homogeneous Maxwell equations

$$\nabla \times \mathbf{E}(\mathbf{r}) = ikZ\mathbf{H}(\mathbf{r}), \nabla \times \mathbf{H}(\mathbf{r}) = -ikY\mathbf{E}(\mathbf{r}) \tag{2.4}$$

except at source points, if any, in the finite part of the plane, it is readily verified that $(\mathbf{E}^0(\mathbf{r}), \mathbf{H}^0(\mathbf{r}))$ also satisfy the homogeneous Maxwell equations except at source points and their images. The quantities $Z$ and $Y$ are respectively the free-space impedance and admittance. Moreover when $z = 0$

$$\hat{\mathbf{z}} \times \mathbf{E}^0(\mathbf{r}) = \mathbf{0}, \qquad \hat{\mathbf{z}} \cdot \mathbf{E}^0(\mathbf{r}) = 2\hat{\mathbf{z}} \cdot \mathbf{E}^{\text{inc}}(\mathbf{r}) \tag{2.5}$$

and

$$\hat{\mathbf{z}} \times \mathbf{H}^0(\mathbf{r}) = 2\hat{\mathbf{z}} \times \mathbf{H}^{\text{inc}}(\mathbf{r}) \quad \hat{\mathbf{z}} \cdot \mathbf{H}^0(\mathbf{r}) = 0. \tag{2.6}$$

Here $\hat{\mathbf{z}}$ denotes a unit vector in the $z$ direction.

The presence of the indentation gives rise to fields in $D$ as well as a scattered field in the upper half space leading to the natural decomposition of the total fields

$$\begin{aligned}\mathbf{E}(\mathbf{r}) &= \mathbf{E}^0(\mathbf{r}) + \mathbf{E}^s(\mathbf{r})\\ \mathbf{H}(\mathbf{r}) &= \mathbf{H}^0(\mathbf{r}) + \mathbf{H}^s(\mathbf{r}), \quad z > 0\end{aligned} \tag{2.7}$$

where $(\mathbf{E}^s, \mathbf{H}^s)$ satisfy the Silver-Müller radiation condition for $z \geq 0$. A more precise statement of the scattering problem then is: for a prescribed $(\mathbf{E}^{\text{inc}}, \mathbf{H}^{\text{inc}})$ find $(\mathbf{E}, \mathbf{H})$ in the scattering domain consisting of the indentation $D$, the upper half space, and $\sigma$ such that $(\mathbf{E}, \mathbf{H})$ and $(\mathbf{E}^s, \mathbf{H}^s)$ satisfy Maxwell's equations in $D$ and the upper half space respectively and $\hat{\mathbf{n}} \times \mathbf{E} = 0$ on $\sigma^c$ and $S$ or equivalently

$$\hat{\mathbf{n}} \times \mathbf{E}^s = \mathbf{0} \text{ on } \sigma^c \quad \text{and} \quad \hat{\mathbf{n}} \times \mathbf{E} = 0 \quad \text{on} \quad S. \tag{2.8}$$

In addition we require that $\int_V (|\mathbf{E}|^2 + |\mathbf{H}|^2) dv < \infty$ where $V$ is any bounded subset of the scattering domain. This finite energy condition ensures fulfillment of the edge condition at the intersection of $S$ with the plane. Note that since $\sigma$ is in the scattering domain, $\mathbf{E}$ and $\mathbf{H}$ (and their derivatives) are continuous there.

In this paper we present boundary integral equations over finite boundaries whose solution gives rise through a representation theorem to the solution of the problem described above, which we refer to hereafter as the *indented screen problem*.

## III. Main Results

We define the free-space scalar Green's function as

$$G(\mathbf{r},\mathbf{r}') := -\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} \tag{3.1}$$

and introduce the Dirichlet and Neumann functions for the full plane

$$G_D(\mathbf{r},\mathbf{r}') := G(\mathbf{r},\mathbf{r}') - G(\mathbf{r},\mathbf{r}_i') \tag{3.2}$$

$$G_N(\mathbf{r},\mathbf{r}') := G(\mathbf{r},\mathbf{r}') + G(\mathbf{r},\mathbf{r}_i'). \tag{3.3}$$

Here $\mathbf{r}_i'$ is the image of the point $\mathbf{r}'$ in the $x$-$y$ plane and $G_D$ and $G_N$ satisfy the conditions

$$G_D(\mathbf{r},\mathbf{r}') = \frac{\partial}{\partial z}G_N(\mathbf{r},\mathbf{r}') = 0 \quad \text{for} \quad z = 0. \tag{3.4}$$

We also introduce three dyadic Green's functions

$$\underline{\Gamma}(\mathbf{r},\mathbf{r}') = ik\nabla \times G(\mathbf{r},\mathbf{r}')\underline{\mathbf{I}} \tag{3.5}$$

$$\underline{\Gamma}_1(\mathbf{r},\mathbf{r}') = ik\nabla \times [G_N(\mathbf{r},\mathbf{r}')\underline{\mathbf{I}}_t + G_D(\mathbf{r},\mathbf{r}')\hat{\mathbf{z}}\hat{\mathbf{z}}] \tag{3.6}$$

$$\underline{\Gamma}_2(\mathbf{r},\mathbf{r}') = ik\nabla \times [G_D(\mathbf{r},\mathbf{r}')\underline{\mathbf{I}}_t + G_N(\mathbf{r},\mathbf{r}')\hat{\mathbf{z}}\hat{\mathbf{z}}] \tag{3.7}$$

where $\underline{\mathbf{I}}$ is the identity dyad and $\underline{\mathbf{I}}_t$ is the transverse identity,

$$\underline{\mathbf{I}}_t = \hat{\mathbf{x}}\hat{\mathbf{x}} + \hat{\mathbf{y}}\hat{\mathbf{y}}. \tag{3.8}$$

We take $\nabla$ to operate on $\mathbf{r}$ and $\nabla'$ to operate on $\mathbf{r}'$. These dyadics satisfy the following distributional differential equations

$$\nabla \times \nabla \times \underline{\Gamma} - k^2\underline{\Gamma} = -ik\nabla \times \delta(\mathbf{r},\mathbf{r}')\underline{\mathbf{I}} \tag{3.9}$$

$$\nabla \times \nabla \times \underline{\Gamma}_1 - k^2\underline{\Gamma}_1 = -ik\nabla \times [\delta(\mathbf{r},\mathbf{r}')\underline{\mathbf{I}} + \delta(\mathbf{r},\mathbf{r}_i')(\underline{\mathbf{I}}_t - \hat{\mathbf{z}}\hat{\mathbf{z}})] \tag{3.10}$$

$$\nabla \times \nabla \times \underline{\Gamma}_2 - k^2\underline{\Gamma}_2 = -ik\nabla \times [\delta(\mathbf{r},\mathbf{r}')\underline{\mathbf{I}} - \delta(\mathbf{r},\mathbf{r}_i')(\underline{\mathbf{I}}_t - \hat{\mathbf{z}}\hat{\mathbf{z}})] \tag{3.11}$$

and the boundary conditions

$$\hat{\mathbf{z}} \times \underline{\Gamma}_1 = \underline{0}, \quad \hat{\mathbf{z}} \times \nabla \times \underline{\Gamma}_2 = \underline{0} \quad \text{for} \quad z = 0. \tag{3.12}$$

Now assume that $(\mathbf{E}, \mathbf{H})$ satisfy the scattering problem defined in Section II. Let us introduce the electric current density

$$\mathbf{J}_S := \hat{\mathbf{n}} \times \mathbf{H} \quad \text{on} \quad S \tag{3.13}$$

where $\hat{\mathbf{n}}$ points away from $D$ and the auxiliary current density

$$\mathbf{J}_\sigma := \hat{\mathbf{z}} \times \mathbf{H} \quad \text{on} \quad \sigma. \tag{3.14}$$

Since $S$ is a part of the physical perfectly conducting boundary, $\mathbf{J}_S$ is an actual current density but since $\sigma$ lies in the scattering domain and is not part of the physical boundary, $\mathbf{J}_\sigma$ is a fictitious electric current density. While it is useful to use $\mathbf{J}_\sigma$ in the analysis it is not essential as will be shown. Nevertheless it is in terms of these two currents that we may represent the field according to the following representation theorem:

*If $(\mathbf{E}, \mathbf{H})$ solve the scattering problem defined in Section II, then*

$$\int_\sigma \mathbf{J}_\sigma(\mathbf{r}) \times \nabla G(\mathbf{r},\mathbf{r}')d\sigma + \int_S \mathbf{J}_S(\mathbf{r}) \times \nabla G(\mathbf{r},\mathbf{r}')ds = \frac{1}{2}\mathbf{H}^s(\mathbf{r}'), z' > 0 \tag{3.16}$$

$$2ik\int_\sigma \mathbf{J}_\sigma(\mathbf{r}) \times \nabla G(\mathbf{r},\mathbf{r}')d\sigma + \int_S \mathbf{J}_S(\mathbf{r}) \cdot \underline{\Gamma}_2(\mathbf{r},\mathbf{r}')ds = ik\mathbf{H}(\mathbf{r}'), \mathbf{r}' \in D \tag{3.17}$$

$$\int_S \mathbf{J}_S(\mathbf{r}) \cdot \underline{\Gamma}_1(\mathbf{r},\mathbf{r}')ds = ik\mathbf{H}^s(\mathbf{r}'), \mathbf{r}' \in D_f \tag{3.18}$$

*and*

$$\int_S \mathbf{J}_S(\mathbf{r}) \cdot \nabla \times \underline{\Gamma}_2(\mathbf{r},\mathbf{r}')ds = k^2 Y\mathbf{E}^s(\mathbf{r})', \mathbf{r}' \in D_f. \tag{3.19}$$

All four representations are derived in the Appendix. The first two, (3.16) and (3.17), show that the electromagnetic fields $(\mathbf{E}, \mathbf{H})$ may be represented everywhere in the scattering domain in terms of $\mathbf{J}_S$ on $S$ and $\mathbf{J}_\sigma$ on $\sigma$. For these two current densities we can state the following.

*The current densities in (3.16) and (3.17) are solutions of the coupled pair of boundary integral equations*

$$\int_S \hat{\mathbf{z}} \times [\mathbf{J}_S(\mathbf{r}) \times \nabla G(\mathbf{r},\mathbf{r}')]ds = \mathbf{J}_\sigma(\mathbf{r}') - \frac{1}{2}\mathbf{J}^0(\mathbf{r}'), \mathbf{r}' \in \sigma \tag{3.20}$$

$$\frac{1}{ik}\int_S \hat{\mathbf{n}}' \times [\mathbf{J}_S(\mathbf{r}) \cdot \underline{\Gamma}_2(\mathbf{r},\mathbf{r}')]ds + 2\int_\sigma \hat{\mathbf{n}}' \times [\mathbf{J}_\sigma(\mathbf{r}) \times \nabla G(\mathbf{r},\mathbf{r}')]d\sigma = \frac{1}{2}\mathbf{J}_S(\mathbf{r}'), \mathbf{r}' \in S \tag{3.21}$$

*where*

$$\mathbf{J}^0(\mathbf{r}') = \mathbf{z}' \times \mathbf{H}^0(\mathbf{r}'). \tag{3.22}$$

Equation (3.20) can be considered as a definition of the fictitious current density $\mathbf{J}_\sigma$ in terms of the real one, $\mathbf{J}_S$. We can use this definition in (3.21) to eliminate $\mathbf{J}_\sigma$ altogether and obtain the following.

*The electric current density $\mathbf{J}_S$ in (3.16)–(3.19) is a solution of the boundary integral equation*

$$\frac{1}{2}\mathbf{J}_S(\mathbf{r}') - \hat{\mathbf{n}}' \times [(\mathbf{K} \circ \mathbf{J}_S)(\mathbf{r}')] = \mathbf{F}(\mathbf{r}'), \mathbf{r}' \in S \tag{3.23}$$

*where*

$$\mathbf{F}(\mathbf{r}') := \hat{\mathbf{n}}' \times \int_\sigma \mathbf{J}^0(\mathbf{r}) \times \nabla G(\mathbf{r},\mathbf{r}')d\sigma \tag{3.24}$$

*and*

$$(\mathbf{K} \circ \mathbf{J}_S)(\mathbf{r}') := \frac{1}{ik}\int_S \mathbf{J}_S(\mathbf{r}) \cdot \underline{\Gamma}_2(\mathbf{r},\mathbf{r}')ds + 2\int_S ds \int_\sigma d\sigma''\{\hat{\mathbf{z}} \times [\mathbf{J}_S(\mathbf{r}) \times \nabla G(\mathbf{r},\mathbf{r}'')]\} \times \nabla'' G(\mathbf{r}'',\mathbf{r}') \tag{3.25}$$

In (3.25), the integration variable on $S$ is $\mathbf{r}$, while the one on $\sigma$ is $\mathbf{r}''$ and $\nabla''$ denotes the gradient with respect to $\mathbf{r}''$. The unknown current density can be extricated from the integral over $\sigma$ in (3.25), and we can make the following statement, alternative to the one above.

*The electric current density $\mathbf{J}_S$ in (3.16)–(3.19) is a solution of the boundary integral equation*

$$\frac{1}{2}\mathbf{J}_S(\mathbf{r}') - (\mathbf{L}\circ\mathbf{J}_S)(\mathbf{r}') = \mathbf{F}(\mathbf{r}'), \mathbf{r}' \in S \tag{3.26}$$

*where $\mathbf{F}$ is defined in (3.24), and*

$$(\mathbf{L}\circ\mathbf{J}_S)(\mathbf{r}') := -\int_S ds\mathbf{J}_S(\mathbf{r})\cdot\left\{\frac{1}{ik}\underline{\Gamma}_2(\mathbf{r},\mathbf{r}')\times\hat{\mathbf{n}}' - 2\int_\sigma d\sigma''\nabla G(\mathbf{r},\mathbf{r}'')\times\left[\hat{\mathbf{z}}\times\underline{\mathbf{I}}\frac{\partial G(\mathbf{r}'',\mathbf{r}')}{\partial n'} + (\hat{\mathbf{n}}'\times\hat{\mathbf{z}})\nabla' G(\mathbf{r}'',\mathbf{r}')\right]\right\} \tag{3.27}$$

Once the current density is known, the fields in $D_f$ can be found using representations (3.18) and (3.19). In order to determine the fields in $D$ and its image $D_i$ one must first find $\mathbf{J}_\sigma$ from (3.20) and then use the representations (3.16) and (3.17) to find the scattered magnetic field in $D_i$ and the total magnetic field in $D$. Additional, simpler looking representations for the fields may be obtained in terms of the fictitious magnetic current density

$$\mathbf{M}_\sigma = \hat{\mathbf{z}}\times\mathbf{E} \quad \text{on} \quad \sigma. \tag{3.28}$$

Two examples are

$$2\int_\sigma \mathbf{M}_\sigma(\mathbf{r})\cdot\nabla\times\underline{\Gamma}(\mathbf{r},\mathbf{r}')d\sigma = k^2 Z\mathbf{H}^s(\mathbf{r}'), \quad z' > 0 \tag{3.29}$$

$$2\int_\sigma \mathbf{M}_\sigma(\mathbf{r})\cdot\underline{\Gamma}(\mathbf{r},\mathbf{r}')d\sigma = -ik\mathbf{E}^s(\mathbf{r}'), \quad z' > 0. \tag{3.30}$$

While these are initially appealing because of their simplicity, they involve the calculation of $\mathbf{M}_\sigma(\mathbf{r})$. While it is possible to express this quantity in terms of $\mathbf{J}_S$ and $\mathbf{J}_\sigma$, this calculation involves a number of additional integrations so that the simplicity gained in the representations of the field for $z' > 0$ is paid for by the additional work in finding $\mathbf{M}_\sigma$.

## IV. The Far Field

The far field is most easily determined using the representation (3.18) which we repeat here in expanded form

$$\int_S \mathbf{J}_S(\mathbf{r})\cdot\nabla\times[G_N(\mathbf{r},\mathbf{r}')\underline{\mathbf{I}}_t + G_D(\mathbf{r},\mathbf{r}')\hat{\mathbf{z}}\hat{\mathbf{z}}]ds = \mathbf{H}^s(\mathbf{r}'), \mathbf{r}' \in D_f. \tag{4.1}$$

Recalling the definitions of the Greens functions, (3.2) and (3.3), and employing the standard asymptotic forms for large $r'$ we see that

$$\nabla G_N = \frac{ike^{ikr'}}{4\pi r'}[\hat{\mathbf{r}}' e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}} + \hat{\mathbf{r}}'_i e^{-ik\hat{\mathbf{r}}'_i\cdot\mathbf{r}}] + O\left(\frac{1}{r'^2}\right) \tag{4.2}$$

$$\nabla G_D = \frac{ike^{ikr'}}{4\pi r'}[\hat{\mathbf{r}}' e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}} - \hat{\mathbf{r}}'_i e^{-ik\hat{\mathbf{r}}'_i\cdot\mathbf{r}}] + O\left(\frac{1}{r'^2}\right) \tag{4.3}$$

where $|\mathbf{r}'_i| = |\mathbf{r}'| = r'$ and $\hat{\mathbf{r}}'_i = \mathbf{r}'_i/r'$. Thus in the far field (for $z' > 0$ or $0 \le \theta' < \pi/2$)

$$\mathbf{H}^s(\mathbf{r}') \approx \frac{ike^{ikr'}}{4\pi r'}\int_S\{\mathbf{J}_S(\mathbf{r})\cdot[\hat{\mathbf{r}}' e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}} + \hat{\mathbf{r}}'_i e^{-ik\hat{\mathbf{r}}'_i\cdot\mathbf{r}}]\times\underline{\mathbf{I}}_t + \mathbf{J}_S(\mathbf{r})\cdot[\hat{\mathbf{r}}' e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}} - \hat{\mathbf{r}}'_i e^{-ik\hat{\mathbf{r}}'_i\cdot\mathbf{r}}]\times\hat{\mathbf{z}}\hat{\mathbf{z}}\}ds \tag{4.4}$$

or

$$\mathbf{H}^S(\mathbf{r}') \approx \frac{ike^{ikr'}}{4\pi r'}\int_S\{e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}}[\mathbf{J}_S(\mathbf{r})\times\hat{\mathbf{r}}'\cdot[\underline{\mathbf{I}}_t + \hat{\mathbf{z}}\hat{\mathbf{z}}] + e^{-ik\hat{\mathbf{r}}'_i\cdot\mathbf{r}}[\mathbf{J}_S(\mathbf{r})\times\hat{\mathbf{r}}'_i\cdot[\underline{\mathbf{I}}_t - \hat{\mathbf{z}}\hat{\mathbf{z}}]]\}ds \tag{4.5}$$

This expression may be further simplified by noting that

$$\mathbf{J}_S(\mathbf{r})\times\hat{\mathbf{r}}'\cdot[\underline{\mathbf{I}}_t + \hat{\mathbf{z}}\hat{\mathbf{z}}] = \mathbf{J}_S(\mathbf{r})\times\hat{\mathbf{r}}' \tag{4.6}$$

and

$$\mathbf{J}_S(\mathbf{r})\times\hat{\mathbf{r}}'_i\cdot[\underline{\mathbf{I}}_t - \hat{\mathbf{z}}\hat{\mathbf{z}}] = (\mathbf{J}_S(\mathbf{r})\times\hat{\mathbf{r}}'_i)_i = -(\mathbf{J}_S)_i(\mathbf{r})\times\hat{\mathbf{r}}' \tag{4.7}$$

where

$$(\mathbf{J}_S)_i(\mathbf{r}) = \mathbf{J}_S(\mathbf{r}) - 2\mathbf{J}_S(\mathbf{r})\cdot\hat{\mathbf{z}}\hat{\mathbf{z}}. \tag{4.8}$$

Then (4.5) may be written as

$$\mathbf{H}^s(\mathbf{r}') = \frac{ike^{ikr'}}{4\pi r'}\hat{\mathbf{r}}'\times\int_S[(\mathbf{J}_S)_i(\mathbf{r})e^{-ik\hat{\mathbf{r}}'_i\cdot\mathbf{r}} - \mathbf{J}_S(\mathbf{r})e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}}]ds. \tag{4.9}$$

Note that the term $O\left(\frac{1}{r'^2}\right)$ has been omitted from (4.4), (4.5) and (4.9). The electric far field is easily obtained from the relation

$$\mathbf{E}^s = -Z\hat{\mathbf{r}}'\times\mathbf{H}^s. \tag{4.10}$$

It is observed that the far field is expressed entirely in terms of the current $\mathbf{J}_S$ on the indentation $S$.

## V. The 2D Case

Both the integral representations and integral equations presented previously simplify greatly when it is assumed that the indentation is cylindrical and all field quantities are identical in planes perpendicular to the cylinder axis. Specifically, we assume that the $x$-axis is the symmetry axis, that the scattering surface consists of a curve $S$ in the $y$-$z$ plane and two semi-infinite lines on the $y$ axis, that all field quantities are independent of $x$, and we define the Green's function to be

$$G(\mathbf{r},\mathbf{r}') = -\frac{i}{4}H_0^{(1)}(k|\mathbf{r}-\mathbf{r}'|) \tag{5.1}$$

where $\mathbf{r} = (y,z)$ and $\mathbf{r}' = (y',z')$. The image $\mathbf{r}_i$ of $\mathbf{r}$ is $(y,z)$. We then consider the following two cases.

*A. Transverse Magnetic Polarization:*

Here we assume that

$$\mathbf{E}^{\text{inc}}(\mathbf{r}) = u^{\text{inc}}(\mathbf{r})\hat{\mathbf{x}} = u^{\text{inc}}(y,z)\hat{\mathbf{x}} \tag{5.2}$$

$$\mathbf{E}^0(\mathbf{r}) = u^0(\mathbf{r})\hat{\mathbf{x}}, u^0(\mathbf{r}) = u^{\text{inc}}(\mathbf{r}) - u^{\text{inc}}(\mathbf{r}_i) \tag{5.3}$$

$$\mathbf{E}(\mathbf{r}) = u(\mathbf{r})\hat{\mathbf{x}}, \mathbf{E}^s(\mathbf{r}) = u^s(\mathbf{r})\hat{\mathbf{x}}. \tag{5.4}$$

The boundary condition (2.9) then implies that $u(\mathbf{r}) = 0$ on $\sigma^c \cup S$. Using Maxwell's equations we find that

$$\mathbf{H}(\mathbf{r}) = \frac{1}{ikZ}\nabla u(\mathbf{r}) \times \hat{\mathbf{x}}. \tag{5.5}$$

The surface current densities become

$$\mathbf{J}_S = -\frac{1}{ikZ}\frac{\partial u}{\partial n}\hat{\mathbf{x}} \quad \text{on} \quad S \tag{5.6}$$

$$\mathbf{J}_\sigma = -\frac{1}{ikZ}\frac{\partial u}{\partial z}\hat{\mathbf{x}} \quad \text{on} \quad \sigma \tag{5.7}$$

and

$$\mathbf{M}_\sigma = u\hat{\mathbf{y}} \quad \text{on} \quad \sigma \tag{5.8}$$

The basic "current density" is $\frac{\partial u}{\partial n}$ on $S$ while $u$ and $\frac{\partial u}{\partial n}$ on the line segment $\sigma$ are auxiliary "current densities". Operating with the curl on the representations in (3.16) and (3.17), and taking into account (5.1)–(5.8), leads to the representations

$$2\int_\sigma \frac{\partial u(\mathbf{r})}{\partial z}G(\mathbf{r},\mathbf{r}')d\sigma + 2\int_S \frac{\partial u(\mathbf{r})}{\partial n}G(\mathbf{r},\mathbf{r}')ds = -u^s(\mathbf{r}'), z' > 0 \tag{5.9}$$

and

$$-2\int_\sigma \frac{\partial u(\mathbf{r})}{\partial z}G(\mathbf{r},\mathbf{r}')d\sigma - \int_S \frac{\partial u(\mathbf{r})}{\partial n}G_N(\mathbf{r},\mathbf{r}')ds = u(\mathbf{r}'), \mathbf{r}' \in D \tag{5.10}$$

with the integrations being with respect to arc length. Equation (5.9) may be rewritten using Green's theorem in $D$ and the fact that $u = 0$ on $S$ as

$$2\int_\sigma u(\mathbf{r})\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial z}d\sigma = -u^s(\mathbf{r}), z' > 0 \tag{5.11}$$

which may also be obtained directly from (3.30). Since the single layer distribution is continuous, (5.9) holds for $z' = 0$ also, in which case it defines $u$ (or equivalently $u^s$) on $\sigma$ in terms of $\frac{\partial u}{\partial n}$ on $\sigma \cup S$. Equation (3.19) becomes

$$\int_S \frac{\partial u(\mathbf{r})}{\partial n}G_N(\mathbf{r},\mathbf{r}')ds = u^s(\mathbf{r}'), \mathbf{r}' \in D, \tag{5.12}$$

in agreement with Willers [20]. The integral equations (3.20) and (3.21) become

$$\frac{\partial u(\mathbf{r}')}{\partial z'} - \int_S \frac{\partial u(\mathbf{r})}{\partial n}\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial z}ds = \frac{\partial u^i(\mathbf{r}')}{\partial z'}, \mathbf{r}' \in \sigma \tag{5.13}$$

$$-2\int_\sigma \frac{\partial u(\mathbf{r})}{\partial z}\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'}d\sigma - \int_S \frac{\partial u(\mathbf{r})}{\partial n}\frac{\partial G_N(\mathbf{r},\mathbf{r}')}{\partial n'}ds = \frac{1}{2}\frac{\partial u(\mathbf{r}')}{\partial n'}, \mathbf{r}' \in S. \tag{5.14}$$

The auxiliary unknown $\frac{\partial u}{\partial z}$ on $\sigma$ may be eliminated yielding an integral equation of the second kind for the basic current density $\frac{\partial u}{\partial n}$ on $S$

$$\frac{1}{2}\frac{\partial u(\mathbf{r}')}{\partial n'} + \int_S \frac{\partial u(\mathbf{r})}{\partial n}K(\mathbf{r},\mathbf{r}')ds = -2\int_\sigma \frac{\partial u^{\text{inc}}(\mathbf{r})}{\partial z}\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n'}d\sigma, \quad \mathbf{r}' \in S \tag{5.15}$$

where

$$K(\mathbf{r},\mathbf{r}') = \frac{\partial G_N(\mathbf{r},\mathbf{r}')}{\partial n'} + 2\int_\sigma \frac{\partial G(\mathbf{r}'',\mathbf{r})}{\partial z} \times \frac{\partial G(\mathbf{r}'',\mathbf{r}')}{\partial n'}d\sigma'', \quad \mathbf{r},\mathbf{r}' \in S. \tag{5.16}$$

We see that it is sufficient to solve (5.15) for $\frac{\partial u}{\partial n}$ on $S$. Then $u$ and $\frac{\partial u}{\partial z}$ on $\sigma$ may be found in a sequence from (5.9) and (5.13). Once all of these functions are obtained, the representation formulas (5.9)–(5.11) may be used to determine the field at any point in the scattering domain. It should be noted, however, that if the far field is the quantity of interest it may be determined through (5.12) solely in terms of $\frac{\partial u}{\partial n}$ on $S$, the basic current density, to be

$$u^s(\mathbf{r}') = \frac{e^{ikr' - \frac{3\pi i}{4}}}{2\sqrt{2\pi kr'}}\int_S \frac{\partial u(\mathbf{r})}{\partial n}(e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}} - e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}'})ds + O\left(\frac{1}{r'^{3/2}}\right), r' \to \infty. \tag{5.17}$$

*B. Transverse Electric Polarization*

In this case we assume that

$$\mathbf{H}^{\text{inc}}(\mathbf{r}) = v^{\text{inc}}(\mathbf{r})\hat{\mathbf{x}} = v^{\text{inc}}(y,z)\hat{\mathbf{x}} \tag{5.18}$$

$$\mathbf{H}^0(\mathbf{r}) = v^0(\mathbf{r})\hat{\mathbf{x}}, v^0(\mathbf{r}) = v^{\text{inc}}(\mathbf{r}) + v^{\text{inc}}(\mathbf{r}_i) \tag{5.19}$$

$$\mathbf{H}(\mathbf{r}) = v(\mathbf{r})\hat{\mathbf{x}}, \mathbf{H}^s(\mathbf{r}) = v^s(\mathbf{r})\hat{\mathbf{x}}. \tag{5.20}$$

Maxwell's equations then imply that

$$\mathbf{E}(\mathbf{r}) = -\frac{1}{ikY}\nabla v \times \hat{\mathbf{x}} \tag{5.21}$$

and the boundary condition (2.9) becomes

$$\frac{\partial v}{\partial n} = 0 \quad \text{on} \quad \sigma^c \cup S. \tag{5.22}$$

The surface current densities become

$$\mathbf{J}_S = v\hat{\mathbf{n}} \times \hat{\mathbf{x}} \quad \text{on} \quad S \tag{5.23}$$

$$\mathbf{J}_\sigma = v\hat{\mathbf{z}} \times \hat{\mathbf{x}} \quad \text{on} \quad \sigma \tag{5.24}$$

$$\mathbf{M}_\sigma = \frac{1}{ikY}\frac{\partial v}{\partial z}\hat{\mathbf{x}} \quad \text{on} \quad \sigma. \tag{5.25}$$

The basic "current density" is $v$ on $S$ and the field everywhere in space is expressible in terms of this fundamental quantity. With these expressions and using the $2D$ Green's function, (5.1), in the dyadics (3.5)–(3.7), the representation formulas, (3.16) and (3.17) become

$$2\int_\sigma v(\mathbf{r})\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial z}d\sigma + 2\int_S v(\mathbf{r})\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n}ds = v^s(\mathbf{r}'), z' > 0 \tag{5.26}$$

and

$$2\int_\sigma v(\mathbf{r})\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial z}d\sigma+\int_S v(\mathbf{r})\frac{\partial G_D(\mathbf{r},\mathbf{r}')}{\partial n}ds = v(\mathbf{r}'),\ \mathbf{r}'\in D. \tag{5.27}$$

Equation (5.26) may be rewritten using Green's theorem in $D$ and the boundary condition (5.22) as

$$2\int_\sigma \frac{\partial v(\mathbf{r})}{\partial z}G(\mathbf{r},\mathbf{r}')d\sigma = v^s(\mathbf{r}'),\ z'\geq 0, \tag{5.28}$$

a result which may also be obtained directly from (3.20). Note that the continuity of the single layer ensures that (5.28) holds when $z'=0$ and hence, this equation may be used to define $v^s$ on $\sigma$ in terms of $\frac{\partial v}{\partial z}$ on $\sigma$; however, it is preferable to use $v$ on $\sigma$ and $S$ as the basic unknowns. Equation (5.28) represents $v^s$ for $z'>0$ in terms of $\frac{\partial v}{\partial z}$ on $\sigma$ whereas (5.26) requires $v$ on both $\sigma$ and $S$. Equation (3.18) however, yields

$$\int_S v(\mathbf{r})\frac{\partial G_N(\mathbf{r},\mathbf{r}')}{\partial n}ds = v^s(\mathbf{r}'),\mathbf{r}'\in D_f \tag{5.29}$$

which represents $v^s$, at least in $D_f$, in terms of $v$ only on $S$. The integral equations for $v$ on $S$ and $\sigma$ are found from (3.22) and (3.23) to be

$$v(\mathbf{r}')-\int_S v(\mathbf{r})\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial n}ds = v^{\text{inc}}(\mathbf{r}'),\mathbf{r}'\in\sigma \tag{5.30}$$

and

$$\frac{1}{2}v(\mathbf{r}')-\int_S v(\mathbf{r})\frac{\partial G_D(\mathbf{r},\mathbf{r}')}{\partial n}ds-2\int_\sigma v(\mathbf{r})\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial z}d\sigma = 0,\quad \mathbf{r}'\in S. \tag{5.31}$$

While (5.30) and (5.31) may be considered as a system of integral equations for the two unknown functions $v$ on $S$ and $v$ on $\sigma$, the auxiliary unknown $v$ on $\sigma$ may be eliminated yielding a Fredholm equation of the second kind for $v$ on $S$:

$$\frac{v(\mathbf{r}')}{2}-\int_S v(\mathbf{r})K(\mathbf{r},\mathbf{r}')ds = 2\int_\sigma v^{\text{inc}}(\mathbf{r})\frac{\partial G(\mathbf{r},\mathbf{r}')}{\partial z}d\sigma,\ \mathbf{r}'\in S \tag{5.32}$$

where

$$K(\mathbf{r},\mathbf{r}') = \frac{\partial G_D(\mathbf{r},\mathbf{r}')}{\partial n}+2\int_\sigma \frac{\partial G(\mathbf{r},\mathbf{r}'')}{\partial n}\times\frac{\partial G(\mathbf{r}'',\mathbf{r}')}{\partial z''}d\sigma'',\quad \mathbf{r},\mathbf{r}'\in S \tag{5.33}$$

It is clear that $v$ on $S$ is the basic current density. When it is known then (5.30) may be used to define $v$ on $\sigma$ and once these quantities have been found, the representation formulas (5.26) and (5.27) may be used to define $v$ in the entire scattering domain. However, the far field may be determined solely in terms of $v$ on $S$ from (5.29) to be

$$v^s(\mathbf{r}') = -\frac{ke^{ik\mathbf{r}'-i\frac{\pi}{4}}}{2\sqrt{2\pi k\mathbf{r}'}}\int_S v(\mathbf{r})\{\hat{\mathbf{n}}\cdot\hat{\mathbf{r}}'e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}}+\hat{\mathbf{n}}\cdot\hat{\mathbf{r}}'e^{-ik\hat{\mathbf{r}}'\cdot\mathbf{r}}\}ds+O\left(\frac{1}{r'^{3/2}}\right),\quad r'\to\infty \tag{5.34}$$

## VI. Conclusion

In this paper we have derived new integral equations whose solution, used in conjunction with new integral representations also derived here, give the field scattered by an indentation of arbitrary shape in a perfectly conducting plane screen. Integral equations of the second kind for the unknown electric current density on the wall of the indentation are found and it is shown that the field everywhere in space may be represented in terms of this quantity. Additional representations involving fictitious electric and magnetic current densities on the plane interface between the indentation and free space are presented. It is conjectured that, while alternative integral equations for these fictitious currents have exhibited the familiar problem of non–unique solvability at frequencies corresponding to cavity resonances of the structure bounded by the indentation and the interface between the indentation and free space, the equations presented here are uniquely solvable at all frequencies. The simplifications that result when the structure is cylindrical are presented for both transverse electric and transverse magnetic polarizations. Finally the far field representations in terms of integrals of the electric current density only over the walls of the indentation are given for the general $3D$ case as well as the $2D$ cases for both polarizations.

## Appendix

Here we indicate how the integral representations and equations of Section III are derived. The basic dyadic identity for this purpose is

$$\int_V [\nabla\times(\nabla\times\mathbf{a})\cdot\underline{\mathbf{A}}-\mathbf{a}\cdot\nabla\times\nabla\times\underline{\mathbf{A}}]dv = \int_B \hat{\mathbf{n}}\cdot[\mathbf{a}\times(\nabla\times\underline{\mathbf{A}})+(\nabla\times\mathbf{a})\times\underline{\mathbf{A}}]ds \tag{A.1}$$

where $\mathbf{a}$ and $\underline{\mathbf{A}}$ are twice differentiable vector and dyadic valued functions, respectively, in $V$ and $\mathbf{a}\times(\nabla\times\underline{\mathbf{A}})+(\nabla\times\mathbf{a})\times\underline{\mathbf{A}}$ is continuous in $V\cup B$, $V$ denotes a domain in $IR^3$ with boundary $B$ and $\hat{\mathbf{n}}$ denotes the unit normal on $B$ directed away from $V$. This identity is found, for example, in [21]. If $\hat{\mathbf{x}}_\ell$, $\ell=1,2,3$ denote rectangular unit vectors ($\hat{\mathbf{x}}_1=\hat{\mathbf{x}}$, $\hat{\mathbf{x}}_2=\hat{\mathbf{y}}$, $\hat{\mathbf{x}}_3=\hat{\mathbf{z}}$) we define the dyadics

$$\underline{\mathbf{\Gamma}}^{(\ell)}(\mathbf{r},\mathbf{r}') := -\frac{ik}{4\pi}\left(\nabla\times\frac{e^{ik|\mathbf{r}-\mathbf{r}'|}}{|\mathbf{r}-\mathbf{r}'|}\hat{\mathbf{x}}_\ell\right)\hat{\mathbf{x}}_\ell\quad \ell=1,2,3. \tag{A.2}$$

Identifying $\underline{\mathbf{A}}$ with $\underline{\mathbf{\Gamma}}^{(\ell)}$ and taking $V$ to be the domain interior to $B$ and exterior to a ball of small radius centered at $\mathbf{r}'$, the above identity can be used, letting the radius of the ball go to zero, to obtain

*Theorem A.1* *If* $\nabla\times\nabla\times\mathbf{a}-k^2\mathbf{a}=0$ *in* $V$, *then*

$$\int_B \hat{\mathbf{n}}\cdot[\mathbf{a}(\mathbf{r})\times(\nabla\times\underline{\mathbf{\Gamma}}^{(\ell)}(\mathbf{r},\mathbf{r}'))-(\nabla\times\mathbf{a}(\mathbf{r}))\times\underline{\mathbf{\Gamma}}^{(\ell)}(\mathbf{r},\mathbf{r}')]ds = \begin{cases} ik\hat{\mathbf{x}}_\ell\hat{\mathbf{x}}_\ell\cdot\nabla'\times\mathbf{a}(\mathbf{r}'), & \mathbf{r}'\in V\\ 0, & \mathbf{r}'\notin\bar{V}.\end{cases} \tag{A.3}$$

*where* $\bar{V}:=V\cup B$.

The derivation of this theorem is a standard, though sensitive, process but we note that both terms in the integral over

the boundary of the ball contribute, that is, if $B_\epsilon(\mathbf{r}')$ denotes the boundary of the ball of radius $\epsilon$ and center at $\mathbf{r}'$ then

$$\lim_{\epsilon\to 0}\int_{B_\epsilon(\mathbf{r}')}\hat{\mathbf{n}}\cdot\mathbf{a}\times(\nabla\times\underline{\Gamma}^{(\ell)}(\mathbf{r},\mathbf{r}'))ds = -\frac{ik}{3}\hat{\mathbf{x}}_\ell\hat{\mathbf{x}}_\ell\cdot\nabla'\times\mathbf{a}(\mathbf{r}'),\ \mathbf{r}'\in V \tag{A.4}$$

and

$$\lim_{\epsilon\to 0}\int_{B_\epsilon(\mathbf{r}')}\hat{\mathbf{n}}\cdot(\nabla\times\mathbf{a}(\mathbf{r}))\times\underline{\Gamma}^{(\ell)}(\mathbf{r},\mathbf{r}')ds = -\frac{2ik}{3}\hat{\mathbf{x}}_\ell\hat{\mathbf{x}}_\ell\cdot\nabla'\times\mathbf{a}(\mathbf{r}'),\ \mathbf{r}'\in V \tag{A.5}$$

where use is made of the expansion

$$\mathbf{a}(\mathbf{r}) = \mathbf{a}(\mathbf{r}') + (\mathbf{r}-\mathbf{r}')\cdot\nabla'\mathbf{a}(\mathbf{r}') + O(\epsilon^2) \tag{A.6}$$

for $\mathbf{r}$ on $B_\epsilon(\mathbf{r}')$. This process does not yield results for $\mathbf{r}'\in B$ since the limit in (A.4) exists only if the integration is over the entire surface of the ball. Theorem A.1 remains valid if $V$ is unbounded provided that $\mathbf{a}$ satisfies the Silver–Müller radiation condition

$$\lim_{r\to\infty}\mathbf{r}\times\nabla\times\mathbf{a}(\mathbf{r}) + ik\mathbf{r}\mathbf{a}(\mathbf{r}) = 0 \quad \text{uniformly in} \quad \hat{\mathbf{r}}. \tag{A.7}$$

We note that the dyadic Green's function $\underline{\Gamma}(\mathbf{r},\mathbf{r}')$ introduced in (3.5) is given by

$$\underline{\Gamma}(\mathbf{r},\mathbf{r}') = \underline{\Gamma}^{(1)}(\mathbf{r},\mathbf{r}') + \underline{\Gamma}^{(2)}(\mathbf{r},\mathbf{r}') + \underline{\Gamma}^{(3)}(\mathbf{r},\mathbf{r}') \tag{A.8}$$

and, hence, (A.3) may be used to obtain

$$\int_B\hat{\mathbf{n}}\cdot[\mathbf{a}(\mathbf{r})\times(\nabla\times\underline{\Gamma}(\mathbf{r},\mathbf{r}')) + (\nabla\times\mathbf{a}(\mathbf{r}))\times\underline{\Gamma}(\mathbf{r},\mathbf{r}')]ds = ik\nabla'\times\mathbf{a}(\mathbf{r}'),\quad \mathbf{r}'\in V = 0,\quad \mathbf{r}'\notin\overline{V}. \tag{A.9}$$

Identifying $V$ with $D$, the portion of the scattering domain lying below the $x$-$y$ plane in the indented screen problem, and a first with $\mathbf{E}$ and then with $\mathbf{H}$ leads to

*Theorem A.2* *If $\mathbf{E}$ and $\mathbf{H}$ satisfy Maxwell's equations in $D$ and $\hat{\mathbf{n}}\times\mathbf{E} = 0$ on $S$ then*

$$\int_\sigma\hat{\mathbf{n}}\times\mathbf{E}(\mathbf{r})\cdot\nabla\times\underline{\Gamma}(\mathbf{r},\mathbf{r}')d\sigma + ikZ\int_{\sigma\cup S}\hat{\mathbf{n}}\times\mathbf{H}(\mathbf{r})\cdot\underline{\Gamma}(\mathbf{r},\mathbf{r}')ds = 0,\quad z'>0 = -k^2Z\mathbf{H}(\mathbf{r}'),\quad \mathbf{r}'\in D \tag{A.10}$$

*and*

$$\int_{\sigma\cup S}\hat{\mathbf{n}}\times\mathbf{H}(\mathbf{r})\cdot\nabla\times\underline{\Gamma}(\mathbf{r},\mathbf{r}')ds - ikY\int_\sigma\hat{\mathbf{n}}\times\mathbf{E}(\mathbf{r})\cdot\underline{\Gamma}(\mathbf{r},\mathbf{r}')d\sigma = 0,\quad z'>0 = k^2Y\mathbf{E}(\mathbf{r}'),\quad \mathbf{r}'\in D \tag{A.11}$$

*where $\hat{\mathbf{n}} = \hat{\mathbf{z}}$ on $\sigma$. Observe that these representations involve not only the electric current densities on $\sigma$ and $S$ but also the "magnetic current density" on $\sigma$. To obtain representations for $\mathbf{E}^s$ and $\mathbf{H}^s$ when $z'>0$ we proceed as follows.*

By replacing $z'$ by $-z'$ in (A.3) we obtain

$$\int_B\hat{\mathbf{n}}\cdot[\mathbf{a}(\mathbf{r})\times(\nabla\times\underline{\Gamma}^{(\ell)}(\mathbf{r},\mathbf{r}_i')) + (\nabla\times\mathbf{a}(\mathbf{r}))\times\underline{\Gamma}^{(\ell)}(\mathbf{r},\mathbf{r}_i')]ds = ik\hat{\mathbf{x}}_\ell\hat{\mathbf{x}}_\ell\cdot\nabla_i'\times\mathbf{a}(\mathbf{r}_i'),\ \mathbf{r}_i'\in V = 0,\ \mathbf{r}_i'\notin\overline{V} \tag{A.12}$$

where $\nabla_i' = \left(\frac{\partial}{\partial x'},\frac{\partial}{\partial y'},-\frac{\partial}{\partial z'}\right)$ and $\mathbf{r}_i'$ is the image of $\mathbf{r}'$ in the $x$-$y$ plane. We define the dyadics

$$\begin{aligned}\underline{\Gamma}_1(\mathbf{r},\mathbf{r}') &= \underline{\Gamma}^{(1)}(\mathbf{r},\mathbf{r}') + \underline{\Gamma}^{(1)}(\mathbf{r},\mathbf{r}_i') + \underline{\Gamma}^{(2)}(\mathbf{r},\mathbf{r}')\\ &\quad + \underline{\Gamma}^{(2)}(\mathbf{r},\mathbf{r}_i') + \underline{\Gamma}^{(3)}(\mathbf{r},\mathbf{r}') + \underline{\Gamma}^{(3)}(\mathbf{r},\mathbf{r}_i')\\ &= ik\nabla\times\{G_N(\mathbf{r},\mathbf{r}')\hat{\mathbf{x}}\hat{\mathbf{x}} + G_N(\mathbf{r},\mathbf{r}')\hat{\mathbf{y}}\hat{\mathbf{y}} + G_D(\mathbf{r},\mathbf{r}')\hat{\mathbf{z}}\hat{\mathbf{z}}\}\end{aligned} \tag{A.13}$$

with $G_N$ and $G_D$ defined in (3.1)–(3.3). With these definitions it is straightforward to verify that the boundary conditions (3.12) are satisfied; that is,

$$\hat{\mathbf{z}}\times\underline{\Gamma}_1 = \hat{\mathbf{z}}\times\nabla\times\underline{\Gamma}_2 = \underline{0} \quad \text{for} \quad z = 0 \tag{A.14}$$

and moreover that

$$\hat{\mathbf{n}}\cdot\mathbf{a}(\mathbf{r})\times\nabla\times\underline{\Gamma}_1(\mathbf{r},\mathbf{r}') = 2\hat{\mathbf{n}}\cdot\mathbf{a}(\mathbf{r})\times(\nabla\times\underline{\Gamma}(\mathbf{r},\mathbf{r}')) \quad \text{for} \quad z = 0 \tag{A.15}$$

and

$$\hat{\mathbf{n}}\cdot(\nabla\times\mathbf{a}(\mathbf{r}))\times\underline{\Gamma}_2(\mathbf{r},\mathbf{r}') = 2\hat{\mathbf{n}}\cdot(\nabla\times\mathbf{a}(\mathbf{r}))\times\underline{\Gamma}(\mathbf{r},\mathbf{r}') \quad \text{for} \quad z = 0. \tag{A.16}$$

Using the definition of the dyadics $\underline{\Gamma}_1$ and $\underline{\Gamma}_2$ equations (A.3) and (A.12) may be combined to yield the representation

$$\int_B\hat{\mathbf{n}}\cdot[\mathbf{a}(\mathbf{r})\times(\nabla\times\underline{\Gamma}_1(\mathbf{r},\mathbf{r}')) + (\nabla\times\mathbf{a}(\mathbf{r}))\times\underline{\Gamma}_1(\mathbf{r},\mathbf{r}')]ds = ik\nabla'\times\mathbf{a}(\mathbf{r}'),\ \mathbf{r}'\in V,\ \mathbf{r}_i'\in\overline{V} = ik(\nabla_i'\times\mathbf{a}(\mathbf{r}_i')),\ \mathbf{r}'\in\overline{V},\ \mathbf{r}_i'\in V = 0,\ \mathbf{r}'\notin\overline{V},\ \mathbf{r}_i'\in\overline{V}. \tag{A.17}$$

If we choose $V$ to be the entire upper-half space, require that $\mathbf{A}$ satisfy the Silver–Müller radiation conditions in it and take into account the boundary conditions (A.14) satisfied by the dyadics, we see that (A.17) implies

$$-\int_{\sigma\cup\sigma^c}\hat{\mathbf{z}}\cdot\mathbf{a}(\mathbf{r})\times(\nabla\times\underline{\Gamma}_1(\mathbf{r},\mathbf{r}'))d\sigma = ik\nabla'\times\mathbf{a}(\mathbf{r}'),\ z'>0 = ik(\nabla_i'\times\mathbf{a}(\mathbf{r}_i')),\ \mathbf{r}'\in D \tag{A.18}$$

$$-\int_{\sigma\cup\sigma^c}\hat{\mathbf{z}}\cdot(\nabla\times\mathbf{a}(\mathbf{r}))\times\underline{\Gamma}_2(\mathbf{r},\mathbf{r}')d\sigma = ik\nabla'\times\mathbf{a}(\mathbf{r}'),\ z'>0 = -ik(\nabla_i'\times\mathbf{a}(\mathbf{r}_i')),\ \mathbf{r}'\in D \tag{A.19}$$

In particular, choosing $\mathbf{a}$ to be $\mathbf{E}^s$ in (A.18) and $\mathbf{H}^s$ in (A.19) we find, since $\hat{\mathbf{z}}\times\mathbf{E} = 0$ on $\sigma^c$,

$$2\int_\sigma\hat{\mathbf{z}}\times\mathbf{E}(\mathbf{r})\cdot\nabla\times\Gamma(\mathbf{r},\mathbf{r}')d\sigma = k^2Z\mathbf{H}^s(\mathbf{r}'),\ z'>0 = k^2Z\mathbf{H}_i^s(\mathbf{r}_i'),\ \mathbf{r}'\in D \tag{A.20}$$

and

$$2\int_\sigma\hat{\mathbf{z}}\times\mathbf{E}(\mathbf{r})\cdot\underline{\Gamma}(\mathbf{r},\mathbf{r}')d\sigma = -ik\mathbf{E}^s(\mathbf{r}'),\ z'>0 = ik\mathbf{E}_i^s(\mathbf{r}_i'),\ \mathbf{r}'\in D. \tag{A.21}$$

To obtain (A.20) and (A.21) we have used Maxwell's equations and the simplifications in (A.15) and (A.16). Note that we may write $\mathbf{E}$ or $\mathbf{E}^s$ in the integrals since $\hat{\mathbf{z}} \times \mathbf{E}^0 = 0$ on $\sigma$. This establishes the representations (3.29) and (3.30). Combining the representations (A.20) and (A.21) with those obtained earlier, (A.10) and (A.11), to eliminate the terms involving $\hat{\mathbf{z}} \times \mathbf{E}$ we find that

$$ikZ \int_{\sigma \cup S} \hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}) \cdot \underline{\Gamma}(\mathbf{r}, \mathbf{r}') ds = -\frac{k^2}{2} Z \mathbf{H}^s(\mathbf{r}'), z' > 0$$
$$= -k^2 Z \mathbf{H}(\mathbf{r}') - \frac{k^2}{2} Z \mathbf{H}_i^s(\mathbf{r}_i'), \ \mathbf{r}' \in D \quad \text{(A.22)}$$

and

$$\int_{\sigma \cup S} \hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}) \cdot \nabla \times \underline{\Gamma}(\mathbf{r}, \mathbf{r}') ds = \frac{k^2}{2} Y \mathbf{E}^s(\mathbf{r}'), z' > 0$$
$$= k^2 Y \mathbf{E}(\mathbf{r}') - \frac{k^2}{2} Y \mathbf{E}_i^s(\mathbf{r}_i'), \quad \mathbf{r}' \in D. \quad \text{(A.23)}$$

Introducing the definition of $\underline{\Gamma}$ and the currents $\mathbf{J}_\sigma = \hat{\mathbf{z}} \times \mathbf{H}$ on $\sigma$ and $\mathbf{J}_S = \hat{\mathbf{n}} \times \mathbf{H}$ on $S$ we obtain the first of two main representations which we state as

*Theorem A.3*

*If* $(\mathbf{E}, \mathbf{H})$ *solve the indented screen problem, then*

$$\int_\sigma \mathbf{J}_\sigma(\mathbf{r}) \times \nabla G(\mathbf{r}, \mathbf{r}') d\sigma + \int_S \mathbf{J}_S(\mathbf{r}) \times \nabla G(\mathbf{r}, \mathbf{r}') ds$$
$$= \frac{1}{2} \mathbf{H}^s(\mathbf{r}'), z' > 0$$
$$= \mathbf{H}(\mathbf{r}') + \frac{1}{2} \mathbf{H}_i^s(\mathbf{r}_i'), \mathbf{r}' \in D \quad \text{(A.24)}$$

*and*

$$\int_\sigma [\mathbf{J}_\sigma(\mathbf{r}) \cdot \nabla \nabla G(\mathbf{r}, \mathbf{r}') + k^2 \mathbf{J}_\sigma(\mathbf{r}) G(\mathbf{r}, \mathbf{r}')] d\sigma$$
$$+ \int_S [\mathbf{J}_S(\mathbf{r}) \cdot \nabla \nabla G(\mathbf{r}, \mathbf{r}') - k^2 \mathbf{J}_S(\mathbf{r}) G(\mathbf{r}, \mathbf{r}')] ds$$
$$= -\frac{ikY}{2} \mathbf{E}^s(\mathbf{r}'), z' > 0$$
$$= -ikY \mathbf{E}(\mathbf{r}') + \frac{ikY}{2} \mathbf{E}_i^s(\mathbf{r}_i'), \mathbf{r}' \in D. \quad \text{(A.25)}$$

The first part of this theorem establishes the representation (3.16). To obtain the second vital representation choose $\mathbf{a}$ to be $\mathbf{E}$ and $V$ to be $D$ in (A.17) and successively take $\underline{\Gamma}$ to be $\underline{\Gamma}_1$ and then $\underline{\Gamma}_2$ which then yields

*Theorem A.4 If* $(\mathbf{E}, \mathbf{H})$ *satisfy Maxwell's equations in* $D$ *and* $\hat{\mathbf{n}} \times \mathbf{E} = 0$ *on* $S$, *then*

$$2 \int_\sigma (\hat{\mathbf{z}} \times \mathbf{E}(\mathbf{r})) \cdot \nabla \times \underline{\Gamma}(\mathbf{r}, \mathbf{r}') d\sigma + ikZ \int_S \hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}) \cdot \underline{\Gamma}_1(\mathbf{r}, \mathbf{r}') ds$$
$$= -k^2 Z \mathbf{H}(\mathbf{r}'), \mathbf{r}' \in D$$
$$= -k^2 Z \mathbf{H}_i(\mathbf{r}_i'), \mathbf{r}_i' \in D$$
$$= 0, \mathbf{r}' \in D_f \quad \text{(A.26)}$$

*and*

$$2 \int_\sigma \hat{\mathbf{z}} \times \mathbf{H}(\mathbf{r}) \cdot \underline{\Gamma}(\mathbf{r}, \mathbf{r}') d\sigma + \int_S \hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}) \cdot \underline{\Gamma}_2(\mathbf{r}, \mathbf{r}') ds$$
$$= ik \mathbf{H}(\mathbf{r}'), \mathbf{r}' \in D$$
$$= -ik \mathbf{H}_i(\mathbf{r}_i'), \mathbf{r}_i' \in D$$
$$= 0, \mathbf{r}' \in D_f. \quad \text{(A.27)}$$

We recall that $D_f$ is the upper half space excluding the image of $D$ and $S$. In arriving at this theorem we have used the boundary conditions satisfied by $\mathbf{E}$, $\underline{\Gamma}_1$ and $\underline{\Gamma}_2$ as well as (A.15) and (A.16) to simplify expressions containing $\underline{\Gamma}_1$ and $\underline{\Gamma}_2$ on $\sigma$. Using the definition of $\underline{\Gamma}$ we see that (A.27) establishes the representation (3.17). Additional representations may be obtained by choosing $\mathbf{a}$ to be $\mathbf{H}$ in (A.17).

Combining (A.20) with (A.26) to eliminate the term involving $\hat{\mathbf{z}} \times \mathbf{E}$ on $\sigma$ leads to

$$\frac{1}{ik} \int_S (\hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r})) \cdot \underline{\Gamma}_1(\mathbf{r}, \mathbf{r}') ds = \mathbf{H}^s(\mathbf{r}'), \mathbf{r}' \in D_f$$
$$= \mathbf{H}^s(\mathbf{r}') + \mathbf{H}_i(\mathbf{r}_i'), \mathbf{r}_i' \in D$$
$$= \mathbf{H}(\mathbf{r}') + \mathbf{H}_i^s(\mathbf{r}_i'), \mathbf{r}' \in D \quad \text{(A.28)}$$

thus establishing the representation (3.18). Taking the curl of both sides of (A.28) we obtain

$$\int_S \hat{\mathbf{n}} \times \mathbf{H}(\mathbf{r}) \cdot \nabla \times \underline{\Gamma}_2(\mathbf{r}, \mathbf{r}') ds = k^2 Y \mathbf{E}^s(\mathbf{r}'), \mathbf{r}' \in D_f$$
$$= k^2 Y \mathbf{E}^s(\mathbf{r}') - k^2 Y \mathbf{E}_i(\mathbf{r}_i'), \mathbf{r}_i' \in D$$
$$= k^2 Y \mathbf{E}(\mathbf{r}') - k^2 Y \mathbf{E}_i^s(\mathbf{r}_i'), \mathbf{r}' \in D \quad \text{(A.29)}$$

which establishes (3.19). In carrying out the computation leading to (A.29) we used the facts that

$$\frac{\partial G_D}{\partial x'} = -\frac{\partial G_D}{\partial x}, \frac{\partial G_D}{\partial y'} = -\frac{\partial G_D}{\partial y}, \frac{\partial G_D}{\partial z'} = -\frac{\partial G_N}{\partial z}$$

and

$$\frac{\partial G_N}{\partial x'} = -\frac{\partial G_N}{\partial x}, \frac{\partial G_N}{\partial y'} = -\frac{\partial G_N}{\partial y}, \frac{\partial G_N}{\partial z'} = -\frac{\partial G_D}{\partial z} \quad \text{(A.30)}$$

to show that

$$\nabla' \times [\mathbf{a} \cdot \underline{\Gamma}_1(\mathbf{r}, \mathbf{r}')] = \mathbf{A} \cdot \nabla \times \underline{\Gamma}_2(\mathbf{r}, \mathbf{r}') \quad \text{(A.31)}$$

for any vector $\mathbf{A}$ constant with respect to the primed variables.

The right-hand sides of (A.28) and (A.29) are not continuous for $\mathbf{r}'$ on $S_i$, the image of $S$. However, because of the singularities in $\underline{\Gamma}_1$ and $\underline{\Gamma}_2$ the integrals on the left have jumps, not only when $\mathbf{r}' \to S$ but also when $\mathbf{r}' \to S_i$ for then $\mathbf{r}_i' \to S$. It is these jumps that allow us to derive the integral equations. The integral equations follow from a straightforward application of the following

*Theorem A.5* (Müller [22]) *Let* $B$ *denote the smooth boundary of a domain* $V$ *and* $\hat{\mathbf{n}}$ *point away from* $V$. *If* $\mathbf{J}$ *is continuous on* $B$ *then*

$$\lim_{\mathbf{r}' \to B^-} \hat{\mathbf{n}}' \times \int_B \mathbf{J}(\mathbf{r}) \times \nabla G(\mathbf{r}, \mathbf{r}') ds$$
$$= \mp \frac{1}{2} \mathbf{J}(\mathbf{r}) + \hat{\mathbf{n}}' \times \int_B \mathbf{J}(\mathbf{r}) \times \nabla G(\mathbf{r}, \mathbf{r}') ds \quad \text{(A.32)}$$

*where* $\mathbf{r}' \to \mathbf{B}^-$ *means* $\mathbf{r}' \to \mathbf{B}$ *from the exterior of* $\mathbf{V}$ *and* $\mathbf{r}' \to \mathbf{B}^-$ *means* $\mathbf{r}' \to B$ *from the interior of* $\mathbf{V}$.

Applying this result to equation (A.24), letting $z' \to 0$ and recalling that for $z' = 0$,

$$\hat{z} \times \mathbf{H}^s(\mathbf{r}') = \hat{z} \times \mathbf{H}(\mathbf{r}') - \hat{z} \times \mathbf{H}^0(\mathbf{r}') = \mathbf{J}_\sigma(\mathbf{r}') - \mathbf{J}^0(\mathbf{r}'). \tag{A.33}$$

we obtain

$$\begin{aligned} \hat{z} \times \int_\sigma \mathbf{J}_\sigma(\mathbf{r}) \times \nabla G(\mathbf{r}, \mathbf{r}') d\sigma \\ + \hat{z} \times \int_S \mathbf{J}_S(\mathbf{r})' \times \nabla G(\mathbf{r}, \mathbf{r}') ds \\ = \mathbf{J}_\sigma(\mathbf{r}') - \frac{1}{2}\mathbf{J}^0(\mathbf{r}'), \mathbf{r}' \in \sigma. \end{aligned} \tag{A.34}$$

But when $\mathbf{r}$ and $\mathbf{r}'$ are on $\sigma$, then $\frac{\partial G}{\partial z} = 0$ and $\hat{z} \cdot \mathbf{J}_\sigma = \hat{z} \cdot \hat{z} \times \mathbf{H} = 0$; hence, the integral over $\sigma$ vanishes and we arrive at

$$\hat{z} \times \int_S \mathbf{J}_S(\mathbf{r}) \times \nabla G(\mathbf{r}, \mathbf{r}') ds = \mathbf{J}_\sigma(\mathbf{r}') - \frac{1}{2}\mathbf{J}^0(\mathbf{r}'), \mathbf{r}' \in \sigma \tag{A.35}$$

which is equation (3.20).

Finally we apply Theorem A.5 to (A.27) by letting $\mathbf{r}' \to S$. Using the definitions of $\underline{\Gamma}$, $\underline{\Gamma}_2$, $\mathbf{J}_\sigma$ and $\mathbf{J}_S$ we may rewrite (A.27) as

$$\begin{aligned} 2\int_\sigma \mathbf{J}_\sigma(\mathbf{r}) \times \nabla G(\mathbf{r}, \mathbf{r}') d\sigma + \int_S \mathbf{J}_S(\mathbf{r}) \cdot \nabla \\ \times [G(\mathbf{r}, \mathbf{r}')\underline{\mathbf{I}} - G(\mathbf{r}, \mathbf{r}_i')(\underline{\mathbf{I}}_t - \hat{z}\hat{z})] ds \\ = \mathbf{H}(\mathbf{r}'), \quad \mathbf{r}' \in D. \end{aligned} \tag{A.36}$$

The only singularity in the integral over $S$ as $\mathbf{r}' \to S$ is in the term

$$\mathbf{J}_S(\mathbf{r}) \cdot \nabla \times G(\mathbf{r}, \mathbf{r}')\underline{\mathbf{I}} = \mathbf{J}_S(\mathbf{r}) \times \nabla G(\mathbf{r}, \mathbf{r}')$$

so that Theorem A.5 applies directly yielding

$$\begin{aligned} 2\hat{\mathbf{n}}' \times \int_\sigma \mathbf{J}_\sigma(\mathbf{r}) \times \nabla G(\mathbf{r}, \mathbf{r}') d\sigma + \frac{\hat{\mathbf{n}}'}{ik} \times \int_S \mathbf{J}_S(\mathbf{r}) \cdot \underline{\Gamma}_2(\mathbf{r}, \mathbf{r}') ds \\ = \frac{1}{2}\mathbf{J}_S(\mathbf{r}'), \quad \mathbf{r}' \in S \end{aligned} \tag{A.37}$$

thus establishing the validity of (3.21).

## REFERENCES


[1] A. J. Poggio and E. K. Miller, "Integral Equation Solutions of Three-Dimensional Scattering Problems," in *Computer Techniques for Electromagnetics*, R. Mittra, Ed. Oxford, Pergamon Press, 1973.

[2] D. Colton and R. Kress, "Integral Equation Methods in Scattering Theory," New York: Wiley, 1983.

[3] H. Brakhage and P. Werner, "Über das Dirichletsche Aussenraum problem für die Helmholtzsche Schwingungsgleichung," *Arch. Math.* vol. 16, pp. 325–329, 1965.

[4] A. J. Burton and G. F. Miller, "The application of integral equation methods to the numerical solution of some exterior boundary value problems," in *Proc. R. Soc. Lond.*, A. 323, pp. 201–210, 1971.

[5] J. R. Mautz and R. F. Harrington, "A combined-source solution for radiation and scattering from a perfectly conducting body," *IEEE Trans. Antenn. Propagat.* vol. AP- 27, pp. 445–454, 1979.

[6] D. S. Jones, "Integral equations for the exterior acoustic problem," *Q J. Mech. Appl. Math.*, vol. 27, pp. 129–142, 1974.

[7] R. E. Kleinman and G. F. Roach, "On modified Green functions in exterior problems for the Helmholtz equation," in *Proc. R. Soc. Lond.*, A. 383, pp. 313–332, 1982.

[8] G. Jost, "Integral equations with modified fundamental solution in time-harmonic electromagnetic scattering," *IMA J. Appl. Math.*, vol. 40, pp. 129–143, 1988.

[9] A. D. Yaghjian, "Augmented electric and magnetic field integral equations," *Radio Sci.*, vol. 16, pp. 987–1001, 1981.

[10] K.-M. Chen, "A mathematical formulation of the equivalence principle," *IEEE Trans. Microwave Theory Technique*, vol. MTT 37, pp. 1576–1581, 1989.

[11] R. F. Harrington and J. R. Mautz, "A generalized network formulation for aperture problems," *IEEE Trans. Antenn. Propagat.*, vol. AP-24, pp. 870–872, Nov. 1976.

[12] C. H. Liang and D. K. Cheng, "Electromagnetic fields coupled into a cavity with a slot–aperture under resonant conditions", *IEEE Trans. Antenn. Propagat.*, AP-30, pp. 664–672, 1982.

[13] J.-M. Jin and J. L. Volakis, "A finite-element-boundary integral formulation for scattering by three-dimensional cavity-backed apertures," *IEEE Trans. Antenn. Propagat.* AP-39, pp. 97–104, Jan. 1991.

[14] ——, "A hybrid finite element method for scattering and radiation by microstrip patch antennas and arrays residing in a cavity," *IEEE Trans. Antenn. Propagat.* AP-39, pp. 1598–1604, 1991.

[15] S.-K. Jeng, "Scattering from a cavity-backed slit on a ground plane—TE case," *IEEE Trans. Antenn. Propagat.* AP-38, pp. 1523–1529, Oct. 1990.

[16] S.-K. Jeng and S.-T. Tzeng, "Scattering from a cavity-backed slit in a ground plane—TM case," *IEEE Trans. Antenn. Propagat.* AP-39 pp. 661–663, May 1991.

[17] T.-M. Wang and H. Ling, "Electromagnetic scattering from three-dimensional cavities via a connection scheme," *IEEE Trans. Antenn. Propagat.* AP-39, pp. 1505–1513, Oct. 1991.

[18] ——, "A connection algorithm on the problem of EM scattering from arbitrary cavities," *J. Electromagn. Wave Appl.*, vol. 5, no. 3, pp. 301–314, 1991.

[19] A. Willers, "The Helmholtz equation in disturbed half-spaces," *Math. Meth. in the Appl. Sci.*, vol. 9, pp. 312–323, 1987.

[20] ——, "Integraigleichungs methoden bei der Reflexion an lokal gestörten Ebenen", Diss. Univ. Göttingen, 1985.

[21] J. van Bladel, *Electromagnetic Fields*, Hemisphere, Washington, 1985, p. 509.

[22] C. Müller, *Foundations of the Mathematical Theory of Electromagnetic Waves*, Berlin: Springer-Verlag, 1969, p. 205.



**John S. Asvestas** (S'63–M'81–SM'89) was born in Athens, Greece. He received the B.S.E., M.S.E., and Ph.D. from the University of Michigan in 1963, 1965, and 1968, respectively.

He is currently with the Corporate Research Center of Grumman Corporation, where he is engaged in scattering and diffraction studies. Prior to joining Grumman Corporation, he worked at the University of Michigan Radiation Laboratory, the Technical University of Denmark Applied Mathematical Physics Laboratory, the University of Delaware Mathematics Department, and the Radar Systems Group of Hughes Aircraft Company.

**Ralph E. Kleinman** (M'75–SM'83–F'94), photograph and biography not available at the time of publication.

Journal of Computational and Applied Mathematics 42 (1992) 17–35
North-Holland



CAM 1179

# A modified gradient method for two-dimensional problems in tomography *

R.E. Kleinman

*Center for the Mathematics of Waves, Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, United States*

P.M. van den Berg

*Laboratory for Electromagnetic Research, Faculty of Electrical Engineering, Delft University of Technology, Delft, Netherlands*

Received 20 June 1991
Revised 10 September 1991

*Abstract*

Kleinman, R.E. and P.M. van den Berg, A modified gradient method for two-dimensional problems in tomography, Journal of Computational and Applied Mathematics 42 (1992) 17–35.

A method for reconstructing the complex index of refraction of a bounded two-dimensional inhomogeneous object of known geometric configuration from measured scattered field data is presented. This work is an extension of recent results on the direct scattering problem wherein the governing domain integral equation was solved iteratively by a successive over-relaxation technique. The relaxation parameter was chosen to minimize the residual error at each step. Convergence of this process was established for indices of refraction much larger than required for convergence of the Born approximation. For the inverse problem the same technique is applied except in this case both the index of refraction and the field are unknown. Iterative solutions for both unknowns are postulated with two relaxation parameters at each step. They are determined by simultaneously minimizing the residual errors in satisfying the domain integral equation and matching the measured data. This procedure retains the nonlinear relation between the two unknowns. Numerical results are presented for a number of representative two-dimensional objects. The algorithm is shown to be effective in cases where the iterative solution of the direct problem is rapidly convergent.

*Keywords:* Scattering, profile inversion, tomography.

*Correspondence to:* Prof. R.E. Kleinman, Center for the Mathematics of Waves, Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, United States.

* This work was supported under NSF Grant No. DMS-8912593, AFOSR Grant 91-0277 and NATO Grant 0230/88 and a Research Grant from the Stichting Fund for Science, Technology and Research (a companion organization of the Schlumberger Foundation in the United States).

## 1. Introduction

In this paper we show how a novel iterative technique can be used to reconstruct complex indices of refraction of two-dimensional objects from measurements of the field (acoustic or electromagnetic) scattered when the object is illuminated by known sources. The method is an extension of the ideas presented in [9,10]. Essentially the method involves casting the inverse problem as an optimization problem in which the cost functional consists of two terms, one is the defect in matching measured data with the field due to a particular index of refraction and the second is the state equation, an integral equation in which the index of refraction appears and which the field must satisfy. A modified gradient method is employed to solve the optimization problem. By modified gradient we mean that the update of the index of refraction takes place in the direction of the gradient of one term of the cost functional, while the update of the field involves a successive over-relaxation method.

Successive over-relaxation is one of a number of iterative methods for solving operator equations which emerge as special cases of general technique based on least-square error minimization, see [7,11,22]. In a recent paper [8], the application of the over-relaxation method to the integral equation arising in scattering from an inhomogeneous object was presented. There it was shown that the iterative solution of the direct problem converged for much larger indices of refraction than those for which the Born series converged.

The Born approximation or Born series is well known as a tool in attempts to solve inverse problems wherein one tries to determine an unknown index of refraction from measurements of a scattered field on some measurement surface exterior to the scattering object. The essence of this approach involves making an initial guess of the field in the object, the Born approximation, then determining the index of refraction to minimize the discrepancy between the far field and the measured data, next solving the direct problem with this newly determined index of refraction in order to update the field in the object and then determining a new index of refraction to minimize the discrepancy in the far field. This iterative process is continued until the defect in matching the measured data is reduced to an acceptable level. Essentially the updating involves a linearization of the highly nonlinear dependence of the field on the index of refraction. In general there are no rigorous convergence results but the scheme has proven to be of practical utility, see e.g., [2,5,12,17,19].

Our approach, inspired by the success of the over-relaxation method in solving the direct problem, avoids the necessity of solving a direct problem at each step of the iteration. Instead the field update directions are chosen as in the successive over-relaxation method to be the residual error in the integral equation while the index update involves the gradient of the defect. This involves the introduction of two relaxation parameters which must be determined at each step. They are found by simultaneously minimizing the residual errors in the field equation and in matching the measured data. This procedure retains the nonlinear relation between the two unknowns.

In the next section we introduce some notation, formulate the problem more precisely, and present a little more detail on previous approaches to the inverse problem. Section 3 presents a brief summary of the relevant over-relaxation results for the direct problem. The new algorithm for solving the inverse problem is given in Section 4 and the results of some numerical experiments using this algorithm in recovering the index of refraction of a two-dimensional object are presented in Section 5. These results are promising in that they successfully reconstruct indices of refraction of fairly general shape.

## 2. Notation and problem statement

Let $D$ denote the interior of a bounded domain in $\mathbb{R}^2$, with piecewise smooth boundary. A precise mathematical characterization of the assumed smoothness is given in [8]. Erect a Cartesian coordinate system with origin in $D$ and denote points in $\mathbb{R}^2$ as $p=(x_p, y_p)$ and $q=(x_q, y_q)$. The subscripts will be omitted when there is no danger of confusion.

We assume that the penetrable inhomogeneous object $D$ is irradiated successively by a number of known incident fields $u_i^{\text{inc}}$, $i=1,\ldots,I$. For each excitation, the direct scattering problem is modelled by the following transmission problem. For a given incident field $u_i^{\text{inc}}(p)$ determine $u_i$ in $D$ and $u_i^{\text{ext}}$ in ext $D$ (exterior of $D$) such that

$$u_i^{\text{ext}} = u_i^{\text{inc}} + u_i^{\text{sct}}, \tag{1}$$

$$\left[\nabla^2 + k^2 n^2(p)\right] u_i(p) = 0, \quad \text{almost everywhere in } D, \tag{2}$$

$$\left[\nabla^2 + k^2\right] u_i^{\text{sct}}(p) = 0, \quad \text{in ext } D, \tag{3}$$

$$u_i^{\text{ext}} = u_i, \quad \text{on } \partial D, \tag{4}$$

$$\frac{\partial u_i^{\text{ext}}}{\partial \nu} = \frac{\partial u_i}{\partial \nu}, \quad \text{on } \partial D, \tag{5}$$

$$\lim_{r\to\infty} r^{1/2}\left[\frac{\partial u_i^{\text{sct}}}{\partial r} - \mathrm{i}k u_i^{\text{sct}}\right] = 0, \quad \text{uniformly in } \hat{p} = \frac{p}{|p|}, \tag{6}$$

and $u_i$ and $\nabla u_i$ are continuous in $D$, but $\nabla^2 u_i$ may not be if the complex index of refraction $n(p)$ is discontinuous. Here $u_i^{\text{inc}}$ is defined in $\mathbb{R}^2$ and is analytic in $D$, $k$ is assumed constant with $\operatorname{Im}(k) \geq 0$ and the complex index of refraction $n(p)$ is piecewise Hölder continuous in $D$. Further $\partial/\partial\nu$ denotes the derivative in the outward direction normal to $\partial D$, and $r := |p| := \sqrt{x^2+y^2}$.

Introduce the complex contrast $\chi$ by

$$\chi(p) = n^2(p) - 1. \tag{7}$$

Then the direct scattering problem may be reformulated as the domain integral equation

$$u_i(p) = u_i^{\text{inc}}(p) + k^2 \int_D \chi(q) u_i(q) \gamma(p, q)\, \mathrm{d}v_q, \quad p \in D,\ i = 1,\ldots,I, \tag{8}$$

where

$$\gamma(p, q) = \tfrac{1}{2}\mathrm{i} H_0^{(1)}(k|p-q|). \tag{9}$$

If $u_i$ solves (8), then the scattered field is obtained from the representation

$$u_i^{\text{sct}}(p) = k^2 \int_D \chi(q) u_i(q) \gamma(p, q)\, \mathrm{d}vq, \quad p \in \text{ext } D,\ i = 1,\ldots,I. \tag{10}$$

Introduce the operator notation

$$G_{(\chi)} u_i(p) = k^2 \int_D \chi(q) u_i(q) \gamma(p, q)\, \mathrm{d}v_q, \quad p \in D, \tag{11}$$

and

$$L_{(\chi)} u_i = u_i - G_{(\chi)} u_i. \tag{12}$$

If $\chi$ is restricted to lie in $L_\infty(D)$ (which includes piecewise continuous functions), then (8), which is simply

$$L_{(\chi)}u_i(p) = u_i^{\mathrm{inc}}(p), \quad p \in D, \ i = 1, \ldots, I, \tag{13}$$

may be considered as an equation for $u_i(p) \in L^2(D)$ where the norm and inner product are

$$\begin{aligned} \| u_i \|_D &= \left\{ \int_D | u_i(p) |^2 \, \mathrm{d}v_p \right\}^{1/2}, \\ \langle u_{i,1}, u_{i,2} \rangle_D &= \int_D u_{i,1}(p) \bar{u}_{i,2}(p) \, \mathrm{d}v_p. \end{aligned} \tag{14}$$

Assume that $u_i^{\mathrm{sct}}$ is measured on some subset $S \subset \mathrm{ext}\ D$. $S$ may be a surface enclosing $D$ or a set of discrete points exterior to $D$. Define a norm and inner product on $S$ by

$$\begin{aligned} \| g_i \|_S &= \left\{ \int_S | g_i(p) |^2 \, \mathrm{d}s_p \right\}^{1/2}, \quad \text{if } S \text{ is a surface,} \\ &= \left\{ \sum_{j=1}^{J} | g_i(p_j) |^2 \right\}^{1/2}, \quad \text{if } S \text{ consists of } J \text{ discrete points } p_j, \\ \langle g_{i,1}, g_{i,2} \rangle_S &= \int_S g_{i,1}(p) \bar{g}_{i,2}(p) \, \mathrm{d}s_p, \quad \text{if } S \text{ is a surface,} \\ &= \sum_{j=1}^{J} g_{i,1}(p_j) \bar{g}_{i,2}(p_j), \quad \text{if } S = \{p_j\}_{j=1}^{J}. \end{aligned} \tag{15}$$

Denote by $g_i(p)$, $p \in S$, the measured data for each excitation $i$, $i = 1, \ldots, I$, and introduce the operator notation (compare (11))

$$K_{(u_i)}\chi(p) = k^2 \int_D \chi(q) u_i(q) \gamma(p, q) \, \mathrm{d}v_q, \quad p \in S; \tag{16}$$

in what follows it is convenient to distinguish between the operator as a mapping of $\chi u_i$ to $D$ and to $S$, respectively.

The profile inversion problem is that of finding $\chi$ for given $g_i$, or solving the equation

$$K_{(u_i)}\chi(p) = g_i(p), \quad p \in S, \ i = 1, \ldots, I, \tag{17}$$

for $\chi$ subject to the additional condition that $u_i$ and $\chi$ satisfy (13) in $D$. The ill-posed nature of this problem is well known [3]. A frequent approach is to attempt to find $\chi$ and $u_i$ to minimize $\sum_{i=1}^{I} \| g_i - K_{(u_i)}\chi \|_S$. Since $u_i$ depends on $\chi$ through (13) in a highly nonlinear way, most attacks on this problem embody two principles; first a linearization of the nonlinear dependence and second a regularization of the optimization problem. The process is usually carried out iteratively in the following way: if $u_{i,n-1}$ is found, determine $\chi_n$ by minimizing $\sum_{i=1}^{I} \| g_i - K_{(u_{i,n-1})}\chi_n \|_S$ using some kind of regularization and update $u_{i,n-1}$ by solving the equation $L_{(\chi_n)}u_{i,n} = u_i^{\mathrm{inc}}$. The starting value is usually taken to be $u_{i,0} = u_i^{\mathrm{inc}}$ (the Born approximation). This essentially follows the idea of [16] and has been utilized in various forms by many investigators [18,24,25].

Our approach follows this same line of reasoning and incorporates the idea of [25] in using the state equation itself as the regularizer. A novel feature of our approach is that we avoid

solving a forward problem (13) at each step of the iteration by generalizing the successive over-relaxation method [11,22] to solve the direct problem. This avoids the linearization implicit in other approaches [18,24,25].

Specifically we will seek $u_i$ and $\chi$ simultaneously to minimize the functional

$$F = \frac{\sum_{i=1}^{I} \| u_i^{\text{inc}} - L_{(\chi)} u_i \|_D^2}{\sum_{i=1}^{I} \| u_i^{\text{inc}} \|_D^2} + \frac{\sum_{i=1}^{I} \| g_i - K_{(u_i)} \chi \|_S^2}{\sum_{i=1}^{I} \| g_i \|_S^2}. \tag{18}$$

Most approaches treat the regularizer as a penalty term with a coefficient which often must be taken to be very small. We choose to put the two terms in (18) on equal footing and normalize them in the sense that they are both equal to one when $u_i = 0$, $i = 1, \ldots, I$.

The iterative solution of the direct problem is summarized in the next section and provides the motivation of our choice of correction direction for the field in the inversion algorithm.

## 3. The direct problem

Details of a number of iterative procedures for solving the operator equation $L_{(\chi)} u_i = u_i^{\text{inc}}$ are presented in [7,11,22]. For our problem of many excitations they consist of constructing sequences of functions $\{u_{i,n}\}_{n=0}^{\infty}$ and associated residuals $\{r_{i,n}\}_{n=0}^{\infty}$ for each $i$ where

$$r_{i,n} := u_i^{\text{inc}} - L_{(\chi)} u_{i,n}, \quad n \geqslant 0, \quad i = 1, \ldots, I. \tag{19}$$

In the stationary over-relaxation method the sequence of functions $\{u_{i,n}\}$ is defined as follows for each $i$:

$$u_{i,0} \quad \text{arbitrary}, \qquad u_{i,n} = u_{i,n-1} + \alpha r_{i,n-1}, \quad n \geqslant 1, \tag{20}$$

whereas the corresponding successive over-relaxation algorithm is

$$u_{i,0} \quad \text{arbitrary}, \qquad u_{i,n} = u_{i,n-1} + \alpha_n r_{i,n-1}, \quad n \geqslant 1,$$

$$\alpha_n = \frac{\sum_{i=1}^{I} \langle r_{i,n-1}, L_{(\chi)} r_{i,n-1} \rangle_D}{\sum_{i=1}^{I} \| L_{(\chi)} r_{i,n-1} \|_D^2}. \tag{21}$$

The difference in the two methods lies in the relaxation parameter. In the stationary method there is a single, possibly complex, parameter $\alpha$ which must be chosen in some manner while in the successive over-relaxation method there is a new $\alpha_n$ at each step which is completely specified by the requirement that it be chosen to minimize $\sum_{i=1}^{I} \| r_{i,n} \|_D$. It should be noted that what we call stationary and successive over-relaxation methods are simple examples of what [13] calls generalized over-relaxation methods. These are operator analogues of Richardson's iterative method in matrix theory (see, e.g., [23, p.141]) and are descent methods with fixed or variable relaxation parameters or direction coefficients (see, e.g., [6, pp. 61ff.]).

For one excitation it was shown [8] that if $\text{Im}(\chi) \geqslant 0$, $\text{Im}(k) \geqslant 0$ and $\chi$ is piecewise Hölder continuous on $D$, then there exists an $\alpha$ such that (13) may be solved by the stationary

procedure. That is, for each $i$, the sequence $u_{i,n}$ generated by (20) converges in $\|\cdot\|_D$, to the solution of (13). No recipe for finding the best choice of $\alpha$ is available but numerical experiments showed that by choosing $\alpha$ to minimize $\|r_{i,1}\|_D$, which leads to the explicit choice $\alpha = \langle r_{i,0}, L_{(\chi)} r_{i,0}\rangle_D / \|L_{(\chi)} r_{i,0}\|_D^2$, resulted in an iterative method with a wider range of convergence than the Born series which is the stationary over-relaxation method with $\alpha = 1$.

While a convergence proof for the successive over-relaxation method is not available for the integral equation under consideration, numerical experiments indicate not only convergence, but also more rapid convergence than in the stationary case. In the stationary case it is shown [8] that it is always possible to choose a relaxation parameter so that the spectral radius of $I - \alpha L_{(\chi)}$ is less than one, so that the iteration converges. Numerical experiments [11] have shown that the successive over-relaxation method clearly converges for a range of contrasts where the Born series diverges and converges faster than the Born series when the latter converges.

The success of the successive over-relaxation method in the direct problem suggests the generalization to the inverse problem described in the next section.

## 4. The inversion algorithm

Here we propose an iterative inversion algorithm which incorporates the ideas of successive over-relaxation with the choice of relaxation parameters determined by minimizing residual error. Of course now there is an unknown function $\chi$ and a vector function $u_i$, $i = 1, \ldots, I$, while two error terms are incorporated in the functional (18). This generalizes the results of [9,10] to multiple sources and higher dimension.

Bearing in mind the fact that the data may consist of a discrete number of measurements from which the unique reconstruction of a completely arbitrary function would be impossible, we recast the problem somewhat. Introduce two families of linearly independent functions $\{\phi_m(q),\ q \in D\}_{m=1}^{M}$ and $\{\psi_j(p),\ p \in S\}_{j=1}^{J}$. Rather than to attempt to reconstruct $\chi$ we limit ourselves to reconstruct the projection of $\chi$ on the linear span of $\{\phi_m\}_{m=1}^{M}$, an approach also used before [24]. Thus we assume

$$x(q) = \sum_{m=1}^{M} \chi_m \phi_m(q). \tag{22}$$

The choice of the functions $\phi_m$ is somewhat arbitrary but with an eye toward an eventual convergence proof, not presented here, the families $\{\phi_m\}_{m=1}^{M}$ should be ultimately dense (as $M \to \infty$) in the space in which the function $\chi$ is sought. Further, $\{\phi_m\}$ should be piecewise Hölder continuous on $D$ in order to be consistent with the assumptions on $\chi$. In addition these functions should be chosen to incorporate any a priori information about $\chi$ that is available. In the absence of any such information the $\phi_m$ may be chosen to be polynomials or finite-element functions. The choice of the functions $\psi_j$ is also arbitrary, although it would be convenient if they were mutually orthogonal on $S$. If the surface $S$ is a circle or sphere, an obvious choice would be circular or spherical harmonics, in which case the expansion coefficients of the data would be $\langle g_i, \psi_j\rangle_S$. In the event that the data are available only at a discrete number of sample points, $p_j$, $j = 1, \ldots, J$, a useful choice is $\psi_j(p) = \delta(p - p_j)$, in which case the inner product $\langle g_i, \psi_j\rangle_S$ is interpreted as the linear functional $\langle g_i, \psi_j\rangle_S = g_i(p_j)$.

We propose to find the projection of $\chi$ on the linear span $\{\phi_m\}_{j=1}^{M}$, which minimizes the error in the projection of $g_i - K_{(u_i)}\chi$ on the linear span of $\{\psi_j\}_{i=1}^{J}$. In order to do this we express functions on $S$ as vectors whose components are projections onto $\psi_j$. First define the coefficients in (22) as an $M$-component vector

$$\boldsymbol{\chi} = (\chi_1, \chi_2, \ldots, \chi_M)^{\mathrm{T}}. \tag{23}$$

Next introduce this vector $\boldsymbol{\chi}$ into (11) with the definition

$$G_{(\chi)}u_i := \sum_{m=1}^{M} \chi_m G_{(\phi_m)}u_i, \tag{24}$$

and in addition

$$L_{(\chi)}u_i := u_i - G_{(\chi)}u_i. \tag{25}$$

Now define $J$-component vectors from functions on $S$ so that the measured data $g_i(p)$ becomes the data vector

$$\boldsymbol{g}_i = \big(\langle g_i, \psi_1\rangle_S, \langle g_i, \psi_2\rangle_S, \ldots, \langle g_i, \psi_J\rangle_S\big)^{\mathrm{T}}, \tag{26}$$

and in addition

$$K_{(u_i)}\boldsymbol{\chi} = \big(\langle K_{(u_i)}\boldsymbol{\chi}\cdot\boldsymbol{\phi}, \psi_1\rangle_S, \langle K_{(u_i)}\boldsymbol{\chi}\cdot\boldsymbol{\phi}, \psi_2\rangle_S, \ldots, \langle K_{(u_i)}\boldsymbol{\chi}\cdot\boldsymbol{\phi}, \psi_J\rangle_S\big)^{\mathrm{T}}, \tag{27}$$

in which $\boldsymbol{\chi}\cdot\boldsymbol{\phi} = \sum_{m=1}^{M}\chi_m\phi_m$. Define the residuals on $D$ and on $S$ as

$$r_i = u_i^{\mathrm{inc}} - L_{(\chi)}u_i, \qquad \boldsymbol{\rho}_i = \boldsymbol{g}_i - K_{(u_i)}\boldsymbol{\chi}. \tag{28}$$

We then propose the iterative construction of sequences $\{u_{i,n}\}$ and $\{\boldsymbol{\chi}_n\}$ as follows:

$$\begin{aligned} &u_{i,0} = u_i^{\mathrm{inc}}, && \boldsymbol{\chi}_0 = \mathbf{0}, \quad \text{(other starting choices may be made)}\\ &u_{i,n} = u_{i,n-1} + \alpha_n r_{i,n-1}, && \boldsymbol{\chi}_n = \boldsymbol{\chi}_{n-1} + \beta_n \boldsymbol{d}_n,\\ &r_{i,n} = u_i^{\mathrm{inc}} - L_{(\chi_n)}u_{i,n}, && \boldsymbol{\rho}_{i,n} = \boldsymbol{g}_i - K_{(u_{i,n})}\boldsymbol{\chi}_n, \end{aligned} \tag{29}$$

where $\alpha_n$ and $\beta_n$ are in general complex constants which are chosen at each step to minimize

$$F_n = \frac{\sum_{i=1}^{I}\|r_{i,n}\|_D^2}{\sum_{i=1}^{I}\|u_i^{\mathrm{inc}}\|_D^2} + \frac{\sum_{i=1}^{I}\|\boldsymbol{\rho}_{i,n}\|_S^2}{\sum_{i=1}^{I}\|\boldsymbol{g}_i\|_S^2}. \tag{30}$$

Here, the residual errors can recursively be written as

$$\begin{aligned} r_{i,n} &= r_{i,n-1} - \alpha_n L_{(\chi_{n-1})}r_{i,n-1} + \beta_n G_{(d_n)}u_{i,n-1} + \alpha_n\beta_n G_{(d_n)}r_{i,n-1},\\ \boldsymbol{\rho}_{i,n} &= \boldsymbol{\rho}_{i,n-1} - \alpha_n K_{(r_{i,n-1})}\boldsymbol{\chi}_{n-1} - \beta_n K_{(u_{i,n-1})}\boldsymbol{d}_n - \alpha_n\beta_n K_{(r_{i,n-1})}\boldsymbol{d}_n. \end{aligned} \tag{31}$$

Note that for each $n$, the vector $\boldsymbol{\chi}_n$ has $M$ components, while $\boldsymbol{\rho}_{i,n}$ and $\boldsymbol{g}_i$ have $J$ components so that by the norm and inner product on $S$ are meant

$$\begin{aligned} \|\boldsymbol{g}_i\|_S^2 &= \sum_{j=1}^{J}|\langle g_i, \psi_j\rangle_S|^2,\\ \langle g_{i,1}, g_{i,2}\rangle_S &= \sum_{j=1}^{J}\langle g_{i,1}, \psi_j\rangle_S\overline{\langle g_{i,2}, \psi_j\rangle}_S. \end{aligned} \tag{32}$$

Implicit in this definition is the assumption that the functions $\psi_j$ are orthogonal on $S$.

In (29) the function $d_n$, which is the updating direction for $\chi_n$, has to be specified. We choose $d_n$ to be the gradient of the error in matching the measured data at the previous, $(n-1)$st, step. Explicitly, treating $\chi$ and $\bar{\chi}$ as independent variables, we define the $M$-component complex-valued vector $d_n$ to be

$$d_n = \left( \frac{\partial}{\partial \bar{\chi}_1}, \frac{\partial}{\partial \bar{\chi}_2}, \ldots, \frac{\partial}{\partial \bar{\chi}_M} \right)^T \sum_{i=1}^{I} \| \rho_i \|_S^2 \Big|_{u_i = u_{i,n-1}, \chi = \chi_{n-1}}. \tag{33}$$

Carrying out the differentiation we find

$$d_n = -\left( \sum_{i=1}^{I} \langle \rho_{i,n-1}, k_{(u_{i,n-1})} \phi_1 \rangle_S, \sum_{i=1}^{I} \langle \rho_{i,n-1}, k_{(u_{i,n-1})} \phi_2 \rangle_S, \right.$$
$$\left. \ldots, \sum_{i=1}^{I} \langle \rho_{i,n-1}, k_{(u_{i,n-1})} \phi_M \rangle_S \right)^T, \tag{34}$$

where the vector $k_{(u_i)} \phi_m$ is defined as

$$k_{(u_i)} \phi_m = \left( \langle K_{(u_i)} \phi_m, \psi_1 \rangle_S, \langle K_{(u_i)} \phi_m, \psi_2 \rangle_S, \ldots, \langle K_{(u_i)} \phi_m, \psi_J \rangle_S \right)^T, \tag{35}$$

and the operator $K_{(u_i)}$ is defined in (16).

The minimization of the quantity $F_n$ of (30), using (31), leads to a nonlinear problem for the variables $\alpha_n$ and $\beta_n$ at each step, which we solve using the Fletcher–Reeves–Polak–Ribiere conjugate gradient method [14] to find values of $\alpha_n$ and $\beta_n$ which produce a (local) minimum. The starting values of $\alpha_n$ and $\beta_n$ are chosen to be equal to zero. Other solutions of this nonlinear algebraic equation have not been investigated.

In the next section we will demonstrate the performance of the present scheme for some representative examples.

## 5. Numerical results

In this section we present the results of a number of numerical examples. In these examples, the domain $D$ is taken to be a square and this square is subdivided into subsquares of equal sizes. In fact, the domain $D$ need not actually be a square. Other shapes can be achieved by choosing the contrast $\chi$ to be zero over portions of the square and this is illustrated in our examples. Hence the inversion algorithm not only reconstructs the index of refraction, but also locates the scatterer within this square. The integrals in the operator expressions are replaced by a summation of the integrals over the subsquares. Over each subsquare the field function and the contrast function are assumed to be constant (the functions $\phi_m$ are pulse functions). Consistent with this approximation, we replace the integration over each subsquare by a polar integration over a circular domain of equal surface area. Then, the integrations over the subdomains can be carried out analytically [15]. The operator expressions containing the $L$ operator or the $G$ operator have a convolution structure and they can then be computed very efficiently with a Fast Fourier Technique [21].

The measurement domain $S$ is taken to be a set of points $\{p_j\}$ equally spaced on a circle circumscribing the square; so the inner products of a function $v$ on $S$ with $\psi_j$ must be interpreted as

$$(v, \psi_j)_S = v(p_j).$$

The sources $u_i^{\text{inc}}$ will be taken to be line sources located at these same points $\{p_j\}$, so that in our examples $I = J$.

For each configuration we first present convergence results for the direct problem using both the successive over-relaxation method and the Born series. We shall compare the R.M.S. error

$$\text{Err}_n = \frac{\sum_{i=1}^{I} \| r_{i,n} \|}{\sum_{i=1}^{I} \| u_i^{\text{inc}} \|} \tag{36}$$

as a function of the number of iterations $n$. These results will be used to support the conjecture that effectiveness of the inversion algorithm depends on the rapidity of convergence of the over-relaxation method for the direct problem and not on the convergence of the Born series.

We then present the results of the inversion algorithm. The measured data were simulated by solving the direct scattering problem with a conjugate gradient method (CGFFT [21]) while imposing an error criterion with an R.M.S. error $\text{Err}_n < 10^{-10}$. The reconstructed contrasts are presented pictorially, and in addition numerical convergence is shown by plotting the profile error

$$\text{Err}_{(\chi_n)} = \frac{\| \chi - \chi_n \|}{\| \chi \|} \tag{37}$$

and the R.M.S. error $F_n^{1/2}$, defined in (30).

*Configuration I*

As first example we consider a square object with dimensions of $\lambda \times \lambda$, where $\lambda = 2\pi/k$ is the free-space wavelength. The contrast profile is given by

$$\chi = \cos\left(\frac{\pi x}{\lambda}\right) \cos\left(\frac{\pi y}{\lambda}\right), \tag{38}$$

where the origin of the coordinate system is at the center of the object. The object is subdivided into $19 \times 19$ subsquares. A surface plot of this profile over the discretized object domain is presented in Fig. 1. Note that the imaginary part of the complex profile is equal to zero. On a circle of diameter $2\lambda$ around the object we locate 10 (line) receivers at equally spaced points, while the object is irradiated by a (line) source that is located successively at each receiver location, hence $I = J = 10$.

We first solve the direct problem for this configuration by using the successive over-relaxation method of (21) and compare the convergence of this method with that of the Born series which is obtained from the stationary over-relaxation method (20) with $\alpha = 1$. The errors are plotted as a function of the number of iterations in Fig. 2. Although the Born series still seems to converge, at a very slow rate, the successive over-relaxation method converges much faster.

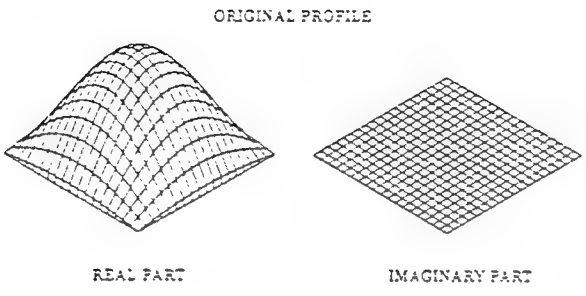



Fig. 1. The contrast profile for Configuration I; the dimensions are $\lambda \times \lambda$; the peak value of the contrast is 1.

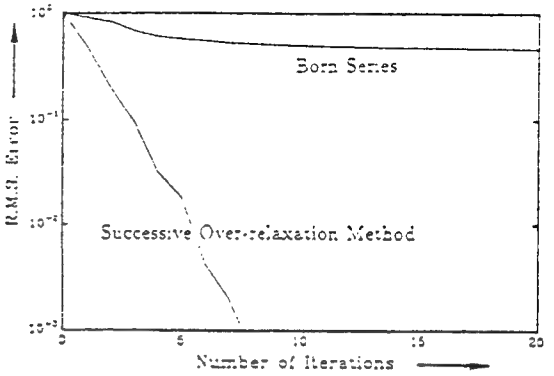



Fig. 2. The numerical convergence of Born series and the successive over-relaxation method in the direct problem for Configuration I.

We secondly solve the inverse problem. The numerical convergence of the profile error $\mathrm{Err}_{(\chi_n)}$ and the R.M.S. error $F_n^{1/2}$ is plotted in Fig. 3. It should be remarked that in practice we are never able to measure the profile error. This means that the error quantity $F_n^{1/2}$ is the only available measure of convergence. We observe that with as few as 15 iterations the errors are decreased to values of about 1%. Some surface plots of the reconstructed profiles are presented in Fig. 4 for various values of $n$, the number of iterations. Comparison with the original profile in Fig. 1 indicates the success of the reconstructions. Note that the imaginary part of the complex profile function converges to zero as it should. Additional numerical experiments confirm that our inversion algorithm reconstructs smooth profiles very accurately as long as the successive over-relaxation method for solving the direct problem converges reasonably rapidly.

*Configuration II*

As second example we consider an object with discontinuous profile. We assume that the object consists of two distinct square homogeneous objects contained inside a square domain with dimensions of $d \times d$. The two objects have diameter of approximately $\frac{1}{4}d$ and the distance between them is also $\frac{1}{4}d$. The contrast or profile function in the larger square has step discontinuities; $\chi = 0$ outside the objects and $\chi = 0.8$ inside the objects. This example is equivalent to that in [1]. However, we use a finer discretization by subdividing the surrounding square into $29 \times 29$ subsquares. A surface plot of this profile over the discretized domain is presented in Fig. 7. Note that the imaginary part of the complex profile is again equal to zero. A circle of diameter $2d$ around the object is equally partitioned by $J$ points. These points serve as receiver locations, while the object is irradiated by a source that is located successively at each receiver location, here $I = J$. We consider three cases, viz. (i) $d = \lambda$ and $I = J = 10$, (ii) $d = 2\lambda$ and $I = J = 20$, (iii) $d = 3\lambda$ and $I = J = 30$.

We first solve the direct problem for this configuration by using the successive over-relaxation method of (21) and compare the convergence of this method with that of the Born series. The numerical results are presented in Fig. 5. We observe that in the case of $d = 3\lambda$ the Born series diverges, while the successive over-relaxation method still converges rapidly.

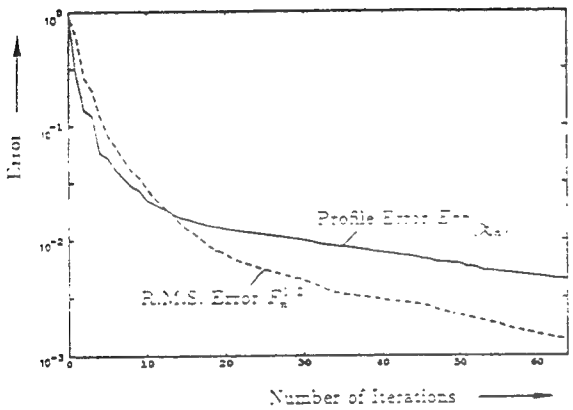



Fig. 3. The numerical convergence of the profile reconstruction for Configuration I.

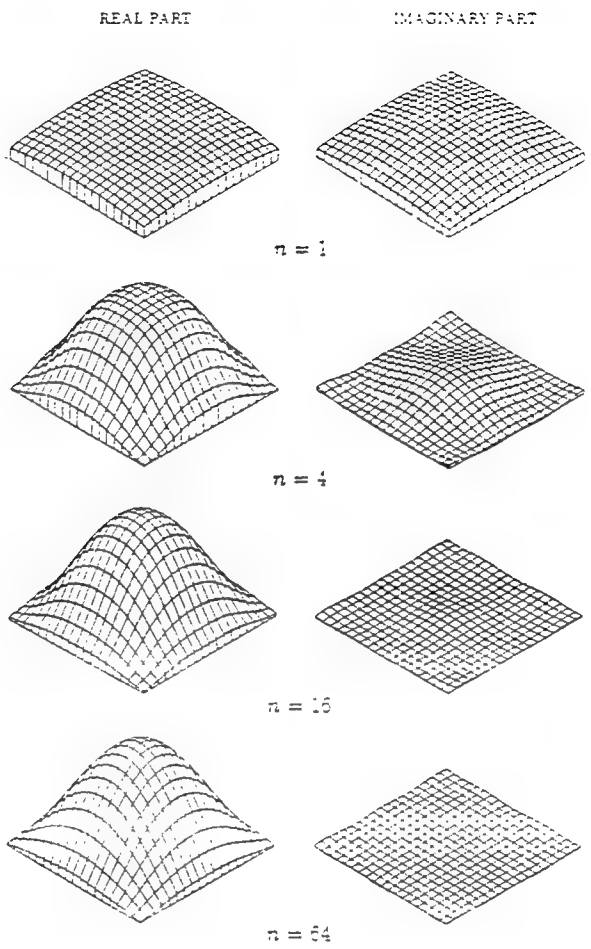



Fig. 4. The reconstructed profiles for Configuration I.

We secondly solve the inverse problem. The numerical convergence of both the profile error $\mathrm{Err}_{(\chi_n)}$ and the R.M.S. error $F_n^{1/2}$ are plotted in Fig. 6. We observe that for the case $d/\lambda = 1$ the profile error remains high as the number of iterations increases. This is also obvious from the surface plots of the reconstructed profiles shown in Fig. 8 for various values of $n$, the number of iterations. It appears that the wavelength of the incident waves is too large to resolve the discontinuities in the profile. Our scheme attempts to reconstruct a band-limited version of the real profile. This observation is in agreement with that of [18, p.310], which states that the expected resolution, using the Rayleigh criterion, is about half a wavelength. We have also performed an additional experiment with 30 transmitters and 30 receivers ($I = J = 30$), but we did not obtain higher resolution. Therefore we have performed some more experiments with smaller wavelengths, viz. $d/\lambda = 2$ and $d/\lambda = 3$. The reconstructed profiles are presented in Figs. 9 and 10, respectively. We indeed see that for decreasing wavelengths a higher resolution is obtained; however, we observe a phenomenon similar to the Gibbs phenomenon in the

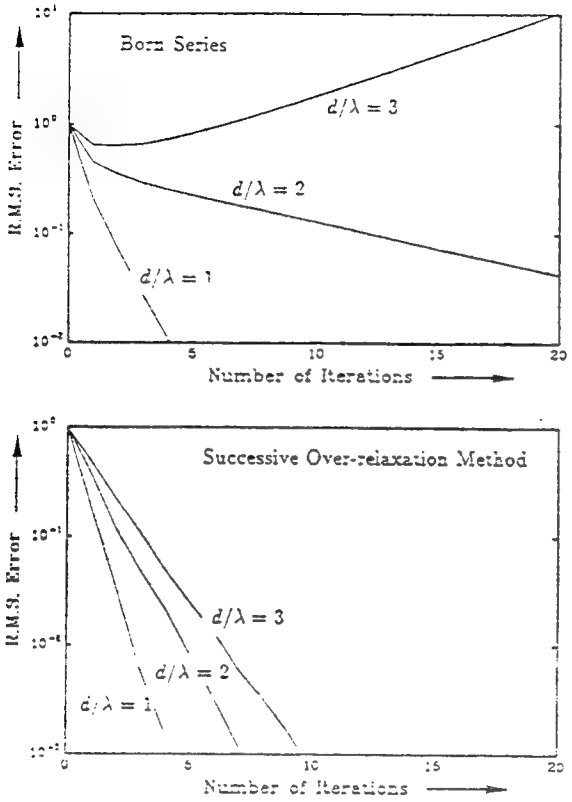



Fig. 5. The numerical convergence of Born series and the successive over-relaxation method in the direct problem for Configuration II, for $d/\lambda = 1, 2$ and 3.

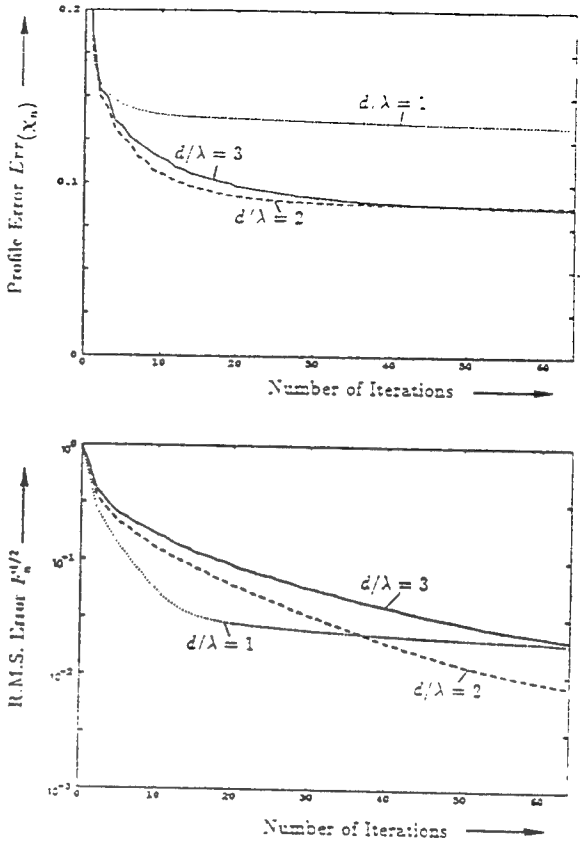



Fig. 6. The numerical convergence of the profile reconstruction for Configuration II, for $d/\lambda = 1, 2$ and 3.

approximation of a discontinuous function by band-limited functions: there occur oscillations near the discontinuities and they increase for smaller wavelengths and they accumulate close to the discontinuities. This confirms that our inversion algorithm is strictly band-limited and the resolution is determined by the wavelength of the incident waves.

*Data with noise*

For our latter example with $d/\lambda = 3$ we investigate the influence of noisy data. We have added to the data a noise signal with maximum amplitude of 10% of the maximum amplitude of the data at all data points $i = 1, \ldots, I$, $j = 1, \ldots, J$. It is observed that this high noise level has only a minor influence on the reconstruction process. In Fig. 11 we observe that the profile error $Err_{(\chi_n)}$ in the case of data with 10% noise behaves almost the same as that without noise, while the R.M.S. error $F_n^{1/2}$ in the case of noisy data is at a very high level (as it should be). Obviously, in the case of noisy data, the R.M.S. error $F_n^{1/2}$ is not a realistic measure of convergence. In Fig. 12 the plots of the reconstructed profiles using noisy data are presented. Comparing Fig. 12 with Fig. 10, we observe the influence of the noise only after large number

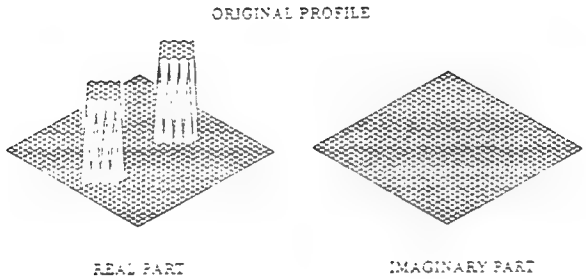



Fig. 7. The contrast profile for Configuration II; the dimensions of the square domain are $d \times d$; the peak value of the contrast is 0.8.

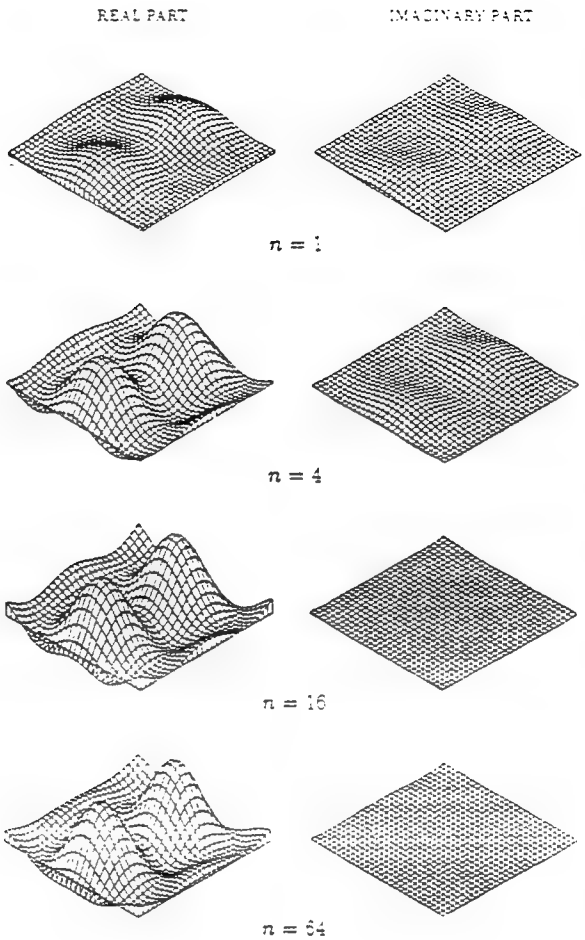



Fig. 8. The reconstructed profiles for Configuration II. when $d = \lambda$.

of iterations. Since the imaginary part of the reconstructed profile has to vanish, the noise is clearly visible in the reconstructed imaginary part of the profile. We believe that the band-limited properties of our inversion scheme make the scheme very robust and not very sensitive to the presence of noise in the data.

*Configuration III*

As last example we consider an object that has a complex contrast with a nonzero imaginary part. We assume the profile function of the object to be defined inside a square domain with dimensions of $3\lambda \times 3\lambda$. The profile distribution is given as follows: inside a square domain of about $\lambda \times \lambda$ the contrast is $\chi_2 = 0.6 + 0.2i$; outside this domain and inside a square domain of about $2\lambda \times 2\lambda$ the contrast is $\chi_1 = 0.3 + 0.4i$; outside the latter domain the contrast vanishes, so that the scattering object is indeed smaller than the square with side $3\lambda$.

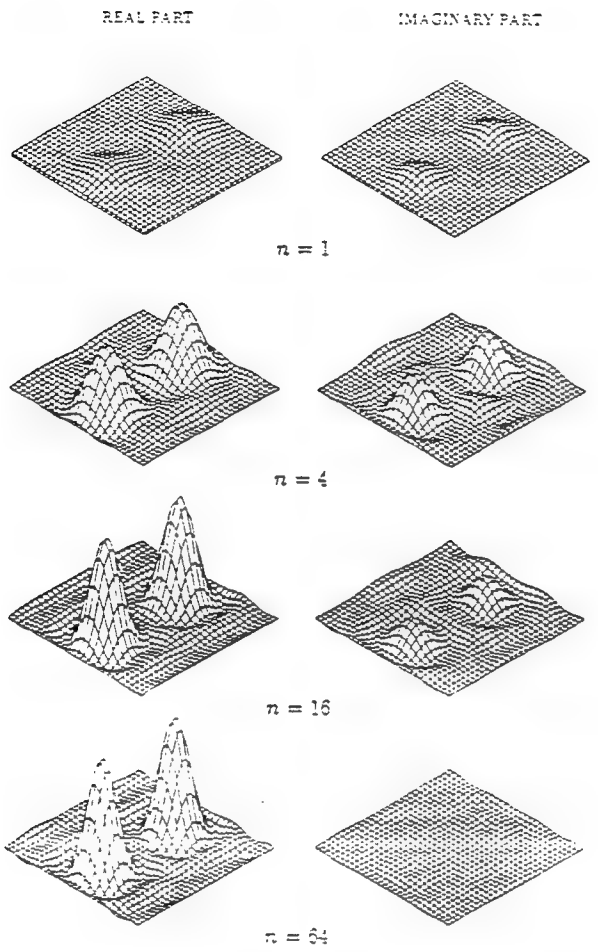



Fig. 9. The reconstructed profiles for Configuration II, when $d = 2\lambda$.

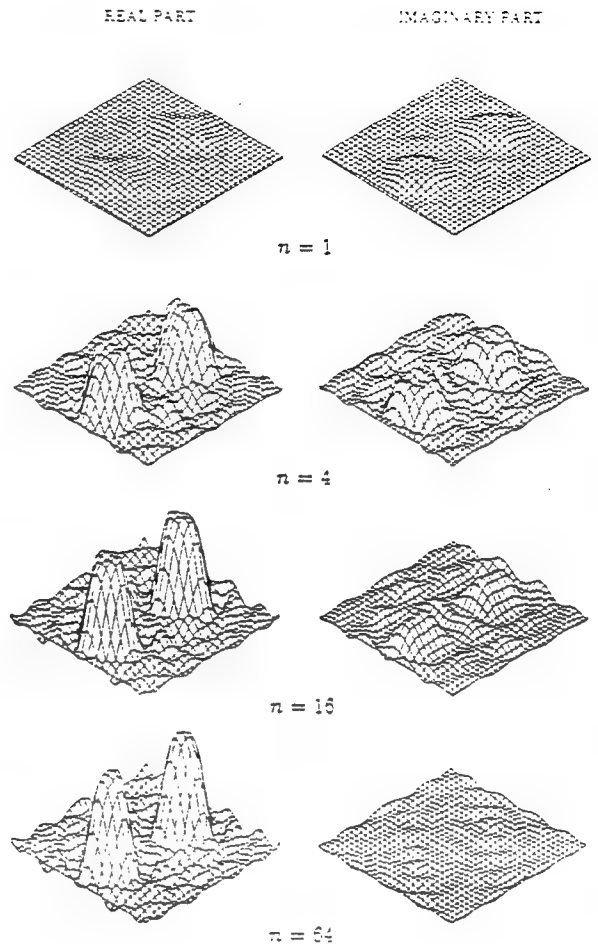



Fig. 10. The reconstructed profiles for Configuration II, when $d = 3\lambda$.

The square is subdivided into $29 \times 29$ subsquares. On a circle of diameter $6\lambda$ around the object 30 receivers are located at equally spaced points, while the object is irradiated by a source that is located successively at each receiver location, hence $I = J = 30$.

As in the previous examples we first solve the direct problem for this configuration by using the successive over-relaxation method of (21) and compare the convergence of this method with that of the Born series. The numerical results are presented in Fig. 13. We observe that the Born series diverges while the successive over-relaxation method converges rapidly.

We secondly solve the inverse problem. The numerical convergence of the profile error $\mathrm{Err}_{(\chi_n)}$ and the R.M.S. error $F_n^{1/2}$ is plotted in Fig. 14 (solid lines) and the reconstructed profiles are given in Fig. 15. A few hundred iterations are needed for reasonable reconstructions. Again it appears that the algorithm tends to reconstruct a band-limited profile, because the profile error remains at a much higher level than the R.M.S. error.

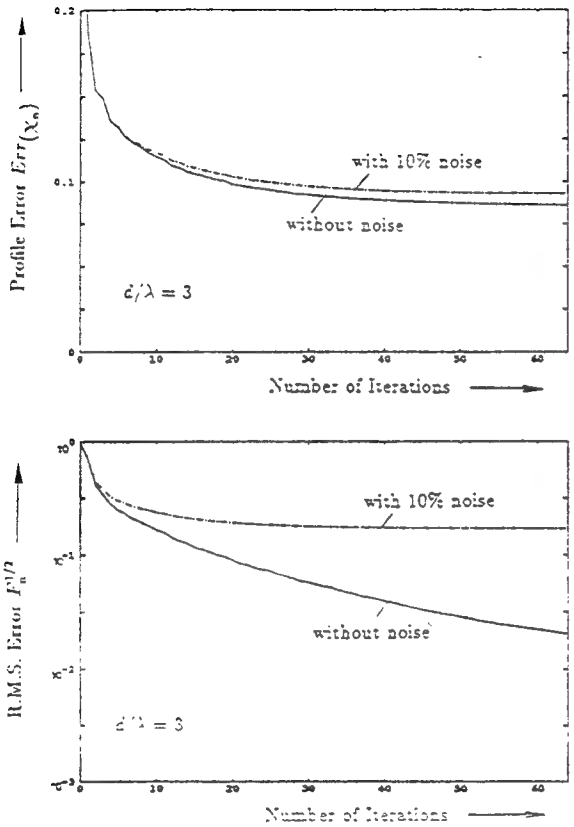



Fig. 11. The numerical convergence of the profile reconstruction for Configuration II, when $d = 3\lambda$, from data with 10% noise and without noise, respectively.

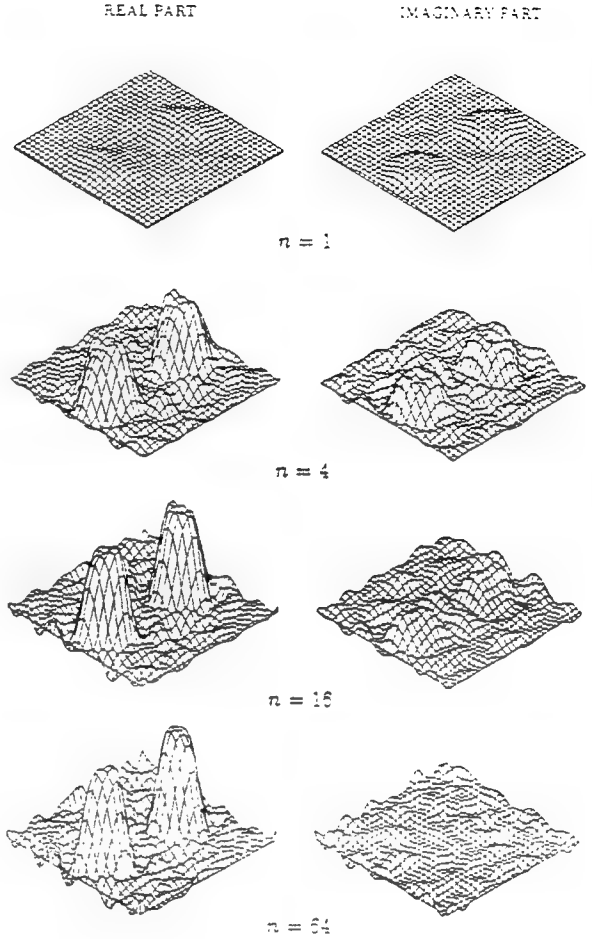



Fig. 12. The reconstructed profiles for Configuration II, when $d = 3\lambda$, from data with 10% noise.

*Band-limited profile*

In order to investigate the phenomenon of band-limitation, we approximate the original profile by a finite Fourier series

$$\chi^{app} = \sum_{j=0}^{7} \sum_{k=0}^{7} \chi_{jk} \cos\left(\frac{j\pi x}{3\lambda}\right) \cos\left(\frac{k\pi y}{3\lambda}\right), \tag{39}$$

where the origin of the coordinate system is at the center of the object. The Fourier coefficients $\chi_{jk}$ are easily determined from the original discontinuous profile shown in Fig. 15. This new profile is taken to be a new original profile and is shown in Fig. 16. Note that this band-limited profile closely resembles the reconstructed profile of Fig. 15 ($n = 512$). Subsequently, we simulated measured data by solving the forward problem for this given band-limited profile. With these data we use our inversion algorithm to reconstruct this profile. The numerical

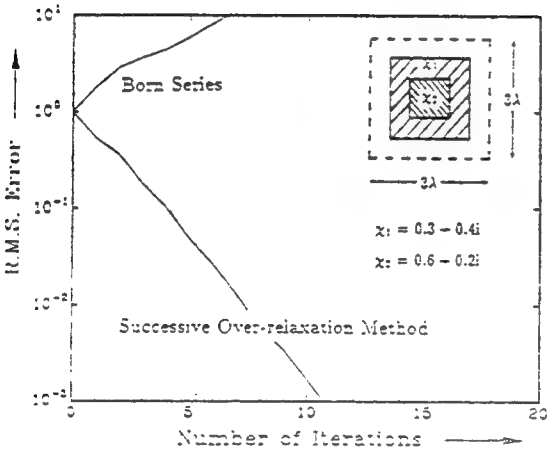



Fig. 13. The numerical convergence of Born series and the successive over-relaxation method in the direct problem for Configuration III (discontinuous profile).

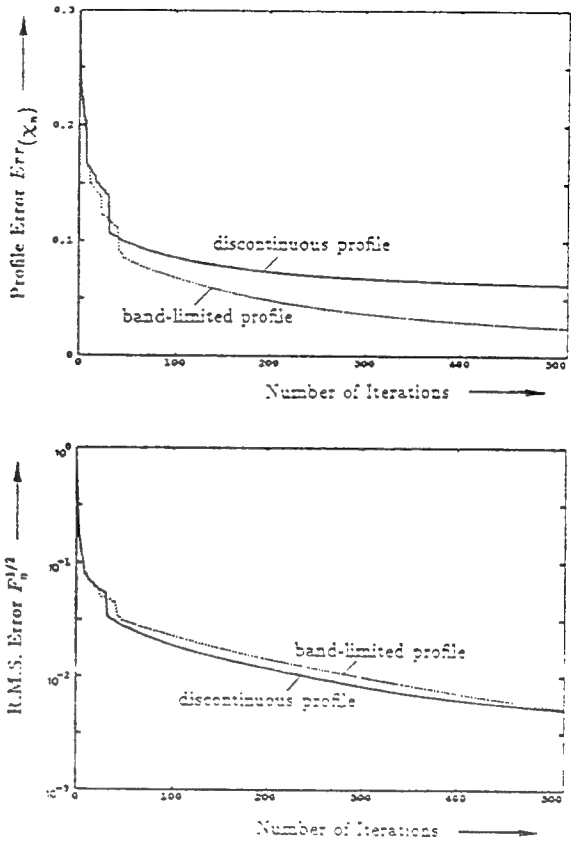



Fig. 14. The numerical convergence of the profile reconstruction for Configuration III.

convergence of the errors are given in Fig. 14 (dotted lines). We observe that the R.M.S. error $F_n^{1/2}$ in the reconstruction of the original discontinuous profile is very close to that of the band-limited profile, but the profile error $Err_{(x_n)}$ in the reconstructed band-limited profile decreases at a much larger rate. The surface plots of the reconstructed profiles for this band-limited case are presented in Fig. 16. Comparing Figs. 15 and 16 we see that the reconstructed profiles are very similar. This supports the assertion that our inversion scheme reconstructs band-limited approximations of the actual profiles.

## 6. Conclusions

In [9,10] we proposed a new iterative scheme to reconstruct the constitutive parameters of a bounded inhomogeneous object from scattering data. The scheme involved the simultaneous minimization of the error in both the object domain and the measurement domain in an iterative way. The method was formulated and tested in one-dimensional problems, scattering

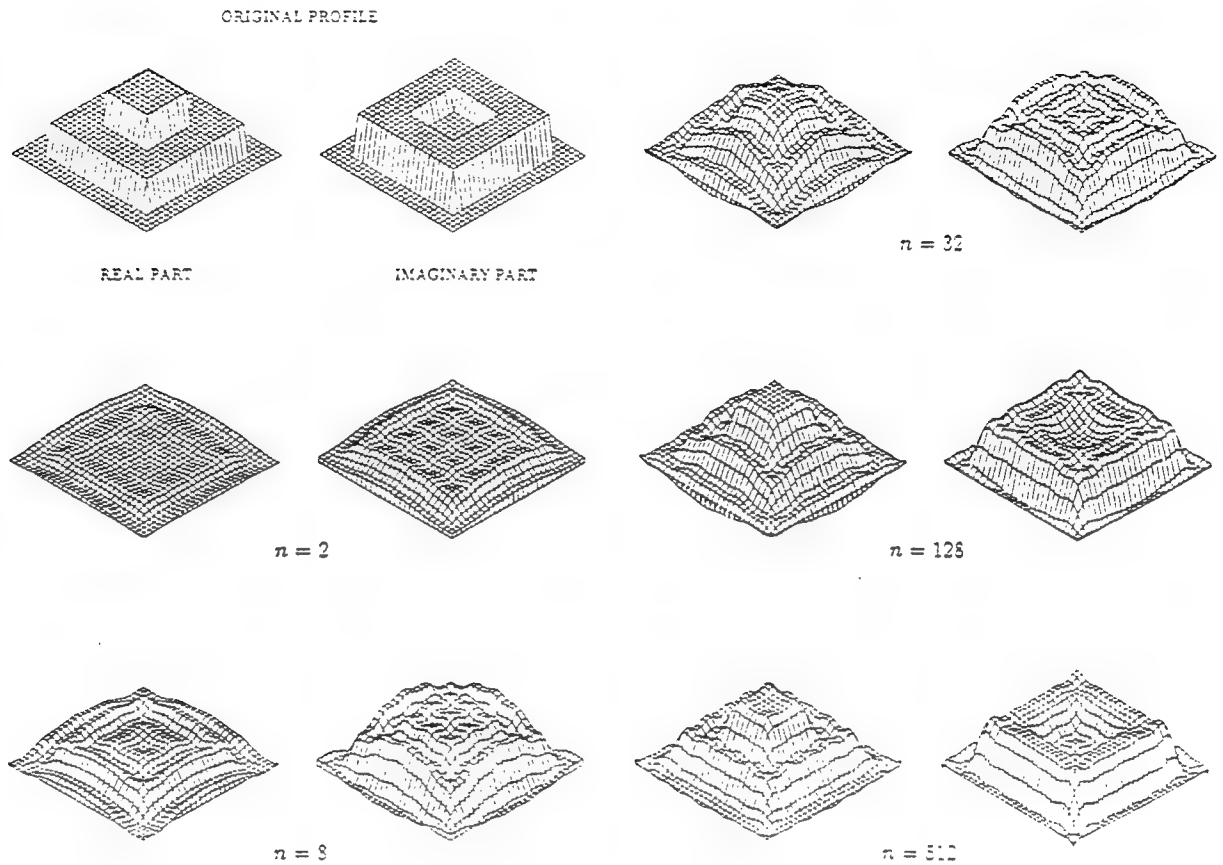



Fig. 15. The reconstructed profiles for Configuration III (discontinuous profile).

by a slab at a single frequency, in which the available data are severely limited (back and forward scattering). This limitation on the number of data points essentially restricted the class of profiles we were able to reconstruct to constant and linear varying indices of refraction.

In the present paper we have extended the algorithm to two dimensions where the number of data points which might be utilized is greatly increased, even at a fixed frequency, by varying source and receiver location. The essential features of the algorithm are retained. It is still based on a successive over-relaxation method for solving the direct problem coupled with a gradient scheme for minimizing the error in matching measured data. While the algorithm is more complicated in that a number of different forward problems corresponding to a number of different incident waves are included, the essential features are the same. It still avoids the need for solving any forward problems at any stage of the iteration, instead the accuracy of the reconstructed contrast and the associated field are increased gradually. A number of numerical examples have been presented which indicate that the algorithm is very effective in reconstructing complex-valued spatially varying contrasts in cases where the successive over-relaxation method produces rapidly convergent solutions of the direct problem. The fact that the success of the reconstruction depends on successive over-relaxation rather than convergence of the Born series means that it is applicable to a wider range of contrasts and frequencies than other

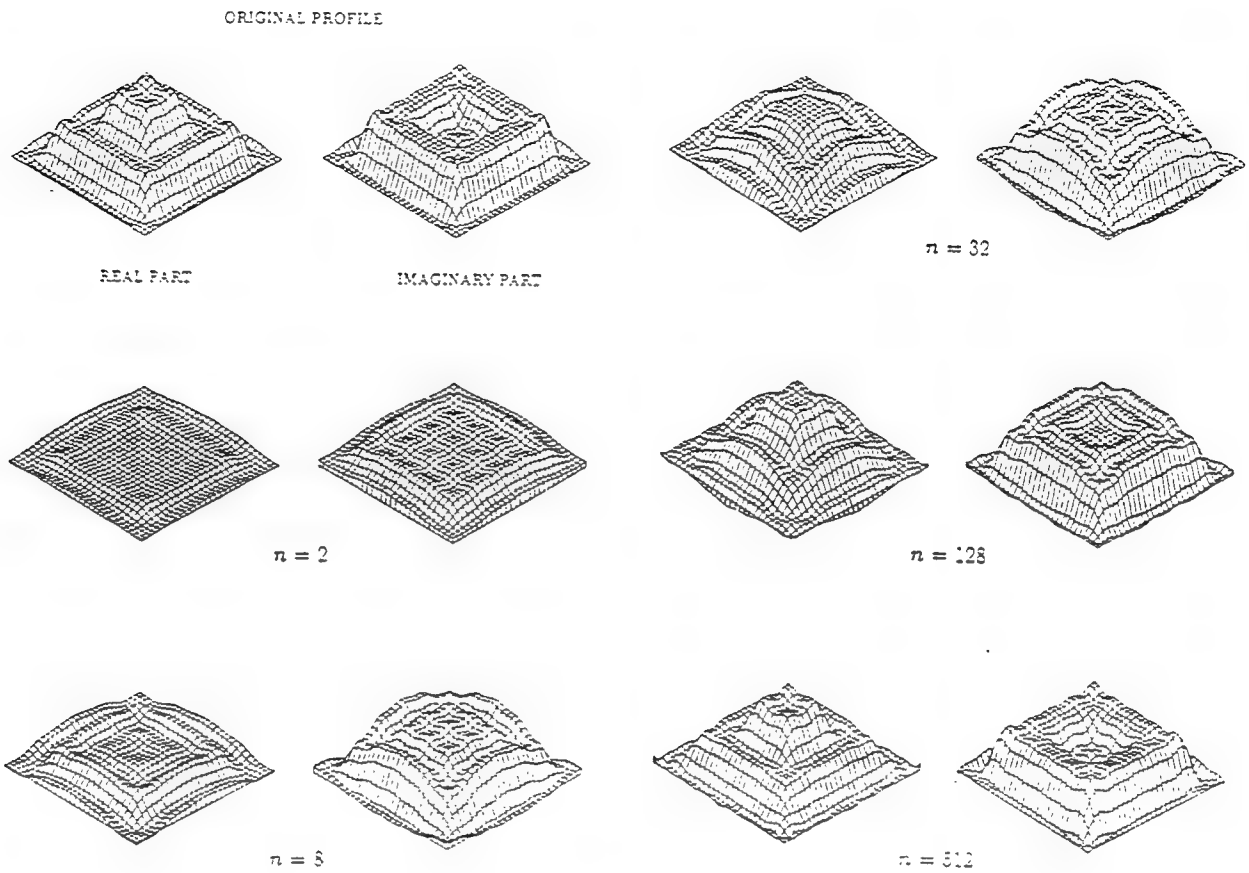



Fig. 16. The reconstructed profiles for Configuration III (band-limited profile).

Born-based inversion methods. The limits on the magnitude of the contrast that can be reconstructed using this method are still to be explored. We expect that a necessary but not sufficient constraint on the contrast is that the successive over-relaxation method must effectively solve the forward or direct problem. However, if a more sophisticated forward solver were incorporated into the algorithm, then even wider ranges contrasts and frequencies could be accommodated, see e.g., [4]. This is one item for future research. Another way to possibly enhance the effectiveness of the method is to build into the scheme the distorted Born iterative method [1,20], but this has yet to be done.

The numerical examples support the contention that spatial variations much less than a wavelength cannot be resolved. Moreover, the way in which the algorithm is constructed, it attempts to reconstruct not the profile itself, but a projection of the profile onto a finite-dimensional space. This effectively imposes a band-limitation on the reconstructed profiles, which is confirmed by the numerical examples.

While we have indicated a number of limitations and possible avenues for future work, the method as it stands appears to constitute an effective tool for profile reconstruction.

## References


[1] W.C. Chew and Y.M. Wang, Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method, *IEEE Trans. Medical Imaging* **9** (1990) 218–225.

[2] A.J. Devaney, A filtered backprojection algorithm for diffraction tomography, *Ultrasonic Imaging* **4** (1982) 336–360.

[3] C.W. Groetsch, *The Theory of Tikhonov Regularization of Fredholm Equations of the First Kind* (Pitman, London, 1984).

[4] T.M. Habashy, E.Y. Chow and D.G. Dudley, Profile inversion using the renormalized source-type integral equation approach, *IEEE Trans. Antennas and Propagation* **AP-38** (1990) 668–682.

[5] T.M. Habashy and R. Mittra, On some inverse methods in electromagnetics, *J. Electromagn. Waves Appl.* **1** (1987) 25–58.

[6] M. Hestenes, *Conjugate Directions in Optimization* (Springer, New York, 1980).

[7] R.E. Kleinman, G.F. Roach, L.S. Schuetz, J. Shirron and P.M. van den Berg, An over-relaxation method for the iterative solution of integral equations in scattering problems, *Wave Motion* **12** (2) (1990) 161–170.

[8] R.E. Kleinman, G.F. Roach and P.M. van den Berg, A convergent Born series for large refractive indices, *J. Opt. Soc. Amer. A* **7** (1990) 890–897.

[9] R.E. Kleinman and P.M. van den Berg, Profile inversion via successive overrelaxation, in: A.F. Gmitro, P.S. Idell and I.J. La Haie, Eds., *Digital Image Synthesis and Inverse Optics*, Proc. SPIE **1351** (1990) 129–139.

[10] R.E. Kleinman and P.M. van den Berg, Non-linearized approach to profile inversion, *Internat. J. Imaging Systems and Technology* **2** (1991) 119–126.

[11] R.E. Kleinman and P.M. van den Berg, Iterative methods for solving integral equations, in: T.K. Sarkar, Ed., *Application of Conjugate Gradient Methods to Electromagnetics and Signal Analysis*, Progr. Electromagn. Research **5** (Elsevier, New York, 1991) Chapter 3.

[12] D. Lesselier, Optimization techniques and inverse problems: Reconstruction of conductivity profiles in the time domain, *IEEE Trans. Antennas and Propagation* AP-30 (1982) 59–65.

[13] W.V. Petryshyn, On a general iterative method for the approximate solution of linear operator equations, *Math. Comp.* **17** (1963) 1–10.

[14] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, *Numerical Recipes. The Art of Scientific Computing* (Cambridge Univ. Press, New York, 1986).

[15] J.H. Richmond, Scattering by a dielectric cylinder of arbitrary cross section shape, *IEEE Trans. Antennas and Propagation* AP-13 (1965) 334–341.

[16] A. Roger, Newton–Kantorovitch algorithm applied to an electromagnetic inverse scattering problem, *IEEE Trans. Antennas and Propagation* **AP-29** (1981) 232–238.

[17] A. Roger, D. Maystre and M. Cadilhac, On a problem of inverse scattering in optics: The dielectric inhomogeneous medium, *J. Optics (Paris)* **9** (1978) 83–90.

[18] W. Tabbara, B. Duchêne, Ch. Pichot, D. Lesselier, L. Chommeloux and N. Joachimowicz, Diffraction tomography: Contribution to the analysis of applications in microwaves and ultrasonics, *Inverse Problems* **4** (1988) 305–331.

[19] A.G. Tijhuis and C. van der Worm, Iterative approach to the frequency-domain scattering solution of the inverse-scattering problem for an inhomogeneous lossless dielectric slab, *IEEE Trans. Antennas and Propagation* AP-32 (1984) 711–716.

[20] W. Tobocman, Iterative inverse scattering method employing Gram–Schmidt orthogonalization, *J. Comput. Phys.* **64** (1986) 230–245.

[21] P.M. van den Berg, Iterative computational techniques in scattering based upon the integrated square error criterion, *IEEE Trans. Antennas and Propagation* **AP-32** (1984) 1063–1071.

[22] P.M. van den Berg and R.E. Kleinman, Iterative solution of integral equations in scattering problems, in: S.K. Datta, J.D. Achenbach and Y.S. Rajapakse, Eds., *Elastic Waves and Ultrasonic Nondestructive Evaluation* (North-Holland, Amsterdam, 1990) 57–62.

[23] R.S. Varga, *Matrix Iterative Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1962).

[24] Y.M. Wang and W.C. Chew, An iterative solution of two-dimensional electromagnetic inverse scattering problem, *J. Imaging Systems* **1** (1989) 100–108.

[25] Z. Xiong, Dreidimensionale geoelektromagnetische Umkehraufgaben, Dissertation, Univ. Göttingen, 1989.

*Radio Science*, Volume 28, Number 5, Pages 877–884, September–October 1993

# An extended range-modified gradient technique for profile inversion

*R. E. Kleinman*

*Center for the Mathematics of Waves, Department of Mathematical Sciences, University of Delaware, Newark*

*P. M. van den Berg*

*Laboratory of Electromagnetic Research, Faculty of Electrical Engineering, Delft University of Technology, Delft, The Netherlands*

(Received November 23, 1992; revised April 7, 1993; accepted April 9, 1993.)

A method for reconstructing the complex index of refraction of a bounded inhomogeneous object from measured scattered field data is presented. The index and the unknown fields within the object are simultaneously reconstructed in an iterative algorithm. The method is a refinement of earlier work which incorporates a more effective way to update the unknowns at each stage of the iteration. Considerable efficiency in the algorithm is achieved. Some numerical examples are given indicating the limits on the contrasts which can be reconstructed. These limits show that the range of contrasts that may be reconstructed is extended over that achievable with the earlier work.

## INTRODUCTION

In previous work [*Kleinman and Van den Berg*, 1992] we presented a novel method for solving the inverse scattering problem of reconstructing the index of refraction of an unknown scatterer from a knowledge of the field scattered when the object is illuminated successively by a number of different excitations. The method was inspired by the success of iterative solutions of the direct scattering problem, and indeed, these iterative methods played a crucial role in the inversion algorithm. The method consists of casting the inverse problem as an optimization problem in which the cost functional is the sum of two terms: one is the defect in matching measured field data with the field scattered by a body with a particular index of refraction, and the second is the error in satisfying the equations of state, a system of integral equations for the field due to each excitation. The index and the fields are updated by a linear iterative method in which the updating directions are weighted by parameters which are determined by minimizing the cost functional. A variety of choices for the updating direction exists, and a relatively simple one has been made by *Kleinman and Van den Berg* [1992] which sufficed to enable some remarkable reconstructions. In the present paper we describe a more sophisticated choice of updating directions which results in a much more efficient algorithm, with an iteration count reduced by approximately a factor of 4. With this more advanced algorithm we analyze the limits of reconstructibility in terms of both object size and the magnitude of refraction index in some two-dimensional examples, and the results of this analysis are presented. Roughly, the upper limit of reconstructibility is found to be $kd|\chi_{max}| = 6\pi$, where $k$ is the wavenumber, $d$ is the object diameter and $|\chi_{max}|$ is the maximum absolute value of the contrast $\chi$. In terms of the index of refraction $n$, $\chi = n^2 - 1$. This problem of reconstruction of the index of refraction has been attached by a number of different methods. The most notable numerical results are given by *Chew and Wang* [1990] and *Colton and Monk* [1992] for real refractive index, and by *Joachimowicz et al.* [1991] and *Habashy et al.* [1992] for complex refractive index. Additional references are given by *Kleinman and Van den Berg* [1992]. Uniqueness for the problem of reconstructing the index of refraction from scattered field data has been proven by *Isakov* [1990] if the scattered field is known in all directions for plane wave incidence from all directions at a single frequency, a situation which is approximated in the present case.

In this paper we will briefly review our inversion algorithm and refer the reader to *Kleinman and Van den Berg* [1992] for more details. Major emphasis will be placed upon the new choice of update



Paper number 93RS01076.
0048-6604/93/93RS-01076$08.00

directions and numerical experiments that probe the limits of applicability of the algorithm.

## DESCRIPTION OF THE TWO-DIMENSIONAL PROBLEM

Assume that a two-dimensional inhomogeneous obstacle $D$ is irradiated successively by a number of known incident fields $u_i^{\text{inc}}$, $i = 1, \cdots, I$. For each excitation the direct scattering problem may be reformulated as the domain integral equation:

$$L_{(\chi)}u_i(\mathbf{p}) = u_i(\mathbf{p}) - G_D\chi u_i(\mathbf{p}) = u_i^{\text{inc}}, \qquad \mathbf{p} \in D, \tag{1}$$

where

$$G_D\chi u_i(\mathbf{p}) = k^2\int_D G(\mathbf{p}, \mathbf{q})\chi(\mathbf{q})u_i(\mathbf{q})\, dv_{\mathbf{q}}, \qquad \mathbf{p} \in D, \tag{2}$$

and

$$G(\mathbf{p}, \mathbf{q}) = \frac{i}{4}H_0^{(1)}(k|\mathbf{p} - \mathbf{q}|). \tag{3}$$

Here $u_i$ is the total field corresponding to the incident field $u_i^{\text{inc}}$, $k$ is the wavenumber, $\chi$ is the complex contrast ($\chi = n^2 - 1$, where $n$ is the index of refraction), and $\mathbf{p}$ and $\mathbf{q}$ are position vectors. $G(\mathbf{p}, \mathbf{q})$ is the free-space Green's function in two dimensions. The inversion algorithm holds equally well in three dimensions with appropriate choice of $G(\mathbf{p}, \mathbf{q})$. $G_D$ is an operator mapping $L^2(D)$ (square integrable functions in $D$) into itself. If $S$ is a surface enclosing $D$, then the scattered field $u_i^{\text{sct}}$ on $S$ is given by $G_S\chi u_i$, where $G_S$ is the same operator defined in (2), except the field point $\mathbf{p}$ now lies on $S$. Hence $G_S$ is an operator mapping $L^2(D)$ into $L^2(S)$. We assume that $u_i^{\text{sct}}$ is measured on $S$ and denote by $f_i(\mathbf{p})$, $\mathbf{p} \in S$, the measured data for each excitation $i$, $i = 1, \cdots, I$. The profile inversion problem is that of finding $\chi$ for given $f_i$, or solving the equations

$$G_S\chi u_i(\mathbf{p}) = f_i(\mathbf{p}), \qquad \mathbf{p} \in S, \qquad i = 1, \cdots, I, \tag{4}$$

for $\chi$, subject to the additional condition that $u_i$ and $\chi$ satisfy (1) in $D$ for each $i$. Thus there are two error measurements involved; the first is the defect in matching the measured data in $L_2(S)$; namely,

$$F_{S,i} = \|f_i - G_S\chi u_i\|_S^2 \tag{5}$$

and the second is the error in the state equation in $L_2(D)$

$$F_{D,i} = \|u_i^{\text{inc}} - L_{(\chi)}u_i\|_D^2, \tag{6}$$

where the subscripts $S$ and $D$ are included in the norm $\|\cdot\|$, and later the inner product $\langle \cdot, \cdot \rangle$ in $L^2$ indicates the domain of integration. Rather than seek $\chi$ to minimize the data error subject to the state error as a constraint, we combine these two error measurements into one normalized cost functional:

$$F = w_D\sum_{i=1}^{I}\|u_i^{\text{inc}} - L_{(\chi)}u_i\|_D^2 + w_S\sum_{i=1}^{I}\|f_i - G_S\chi u_i\|_S^2, \tag{7}$$

where

$$w_D = \left(\sum_{i=1}^{I}\|u_i^{\text{inc}}\|_D^2\right)^{-1} \qquad w_S = \left(\sum_{i=1}^{I}\|f_i\|_S^2\right)^{-1}. \tag{8}$$

If the data $f_i$ originate from an actual scattering problem, then there exists $\chi$ and $u_i$ in $L_2(D)$ for which $F$ vanishes. In practice, we approximate $\chi$ and $u_i$ in subspaces of $L^2(D)$, and hence the guiding principle is to seek $\chi$ and $u_i$ simultaneously in a way which minimizes $F$.

## INVERSION ALGORITHM

The basic idea underlying the inversion algorithm is to incorporate the ideas of a gradient type of algorithm to iteratively solve the direct scattering problem together with a similar algorithm for solving the ill-posed inverse problem. Specifically, we propose the iterative construction of sequences $\{u_{i,n}\}$ and $\{\chi_n\}$ as follows:

$$u_{i,n} = u_{i,n-1} + \alpha_n v_{i,n}, \qquad \chi_n = \chi_{n-1} + \beta_n d_n,$$
$$n = 1, 2 \cdots. \tag{9}$$

The functions $v_{i,n}$ and $d_n$ are update directions for the functions $u_{i,n}$ and $\chi_n$, respectively, while the parameters $\alpha_n$ and $\beta_n$ are weights to be determined. The residual errors at each step in the state equation and data equation are defined as

$$r_{i,n} = u_i^{\text{inc}} - L_{(\chi_n)}u_{i,n}, \qquad \rho_{i,n} = f_i - G_S\chi_n u_{i,n}, \tag{10}$$

and the value of the cost functional at the $n$th step is

$$F_n = w_D\sum_{i=1}^{I}\|r_{i,n}\|_D^2 + w_S\sum_{i=1}^{I}\|\rho_{i,n}\|_S^2. \tag{11}$$

Following *Kleinman and Van den Berg*, [1992] the values of the parameters $\alpha_n$ and $\beta_n$ are determined by requiring $F_n$ to be a minimum. This leads to two nonlinear complexly valued algebraic equations which we write implicitly as

$$w_D \sum_{i=1}^{I} \langle r_{i,n}, -L_{(\chi_{n-1})} v_{i,n} + \beta_n G_D d_n v_{i,n} \rangle_D - w_s \sum_{i=1}^{I} \langle \rho_{i,n}, G_S \chi_{n-1} v_{i,n} + \beta_n G_S d_n v_{i,n} \rangle_S = 0, \quad (12)$$

$$w_D \sum_{i=1}^{I} \langle r_{i,n}, G_D d_n u_{i,n-1} + \alpha_n G_D d_n v_{i,n} \rangle_D - w_s \sum_{i=1}^{I} \langle \rho_{i,n}, G_S d_n u_{i,n-1} + \alpha_n G_S d_n v_{i,n} \rangle_S = 0, \quad (13)$$

where the residuals satisfy the recursive relations

$$r_{i,n} = r_{i,n-1} - \alpha_n L_{(\chi_{n-1})} v_{i,n} + \beta_n G_D d_n u_{i,n-1} + \alpha_n \beta_n G_D d_n v_{i,n}, \quad (14)$$

$$\rho_{i,n} = \rho_{i,n-1} - \alpha_n G_S \chi_{n-1} v_{i,n} - \beta_n G_S d_n u_{i,n-1} - \alpha_n \beta_n G_S d_n v_{i,n}. \quad (15)$$

Substitution of these expressions in (12) and (13) results in two equations involving terms determined at the $(n-1)$st step, the directions $d_n$ and $v_{i,n}$, and the two parameters $\alpha_n$ and $\beta_n$. Once the directions $d_n$ and $v_{i,n}$ are chosen, we have nonlinear algebraic equations in $\alpha_n$ and $\beta_n$, and the solution of these equations is accomplished using the Fletcher-Reeves-Polak-Ribière conjugate gradient method [*Press et al.*, 1986]. The starting value for $\alpha_n$ is obtained by taking $\beta_n = 0$ and minimizing $F_n$, while the starting value for $\beta_n$ is found by setting $\alpha_n = 0$ and again minimizing $F_n$. This procedure retains the nonlinear character of the problem at each step in contrast with other iterative treatments that linearize the problem at every stage [e.g., *Roger*, 1981]. The essential ingredients remaining are the initial choices and the update directions.

## INITIAL GUESS AND CORRECTION DIRECTIONS

In our previous treatment of this problem [*Kleinman and Van den Berg*, 1992] we chose $\chi_0 = 0$ and $u_{i,0} = u_i^{\text{inc}}$. The update direction for the field was directly adapted from the successive overrelaxation method for solving the direct problem with known contrast [*Kleinman and Van den Berg*, 1991] to be $v_{i,n} = r_{i,n-1}$, and the update direction for the contrast was chosen to be the gradient of the error in the measured data at the previous, $(n-1)$st, step. In the present work we refine these choices considerably. As the result of many numerical experiments it was found that substantial advantage could be gained by first reconstructing a best possible constant contrast, even when the contrast in reality was variable. Thus the algorithm was split into two stages, or more precisely, the algorithm was run twice, first to determine the constant $\chi^{\text{initial}}$, using $d_n = 1$, and the associated fields $u_i^{\text{initial}}$, then using these initial values in the algorithm to obtain the final values of $\chi_n$ and $u_{i,n}$. The update directions were chosen in different ways depending on how rapidly corrections were occurring, the idea being that simpler directions should be used when possible. This resulted in significant reductions in computational time.

Specifically, we proceed as follows. Define the normalized change in the field by

$$\varepsilon_n = \left( \sum_{i=1}^{I} \| u_{i,n} - u_{i,n-1} \|_D^2 \right)^{1/2} \left( \sum_{i=1}^{I} \| u_{i,n-1} \|_D^2 \right)^{-1/2}. \quad (16)$$

To determine the initial values, we set an arbitrary switching criterion $\varepsilon$ and run the algorithm of (9) with $\chi_0^{\text{initial}} = 0$, $u_{i,0}^{\text{initial}} = u_i^{\text{inc}}$, $d_n = 1$, and $v_{i,n} = r_{i,n-1}$ until $\varepsilon_{n-1} < \varepsilon$, then switch the definition of $v_{i,n}$ to

$$v_{i,n} = g_{i,n}^v + \gamma_n^v v_{i,n-1} \qquad \gamma_n^v = \frac{\sum_{i=1}^{I} \langle g_{i,n}^v, g_{i,n}^v - g_{i,n-1}^v \rangle_D}{\sum_{i=1}^{I} \| g_{i,n-1}^v \|_D^2}, \quad (17)$$

with the gradient

$$g_{i,n}^v = w_D (r_{i,n-1} - \bar{\chi}_{n-1} \bar{G}_D r_{i,n-1}) + w_S \bar{\chi}_{n-1} \bar{G}_S \rho_{i,n-1}, \quad (18)$$

where the overbar denotes complex conjugate, and $\bar{G}_S$ is a map from $L^2(S)$ to $L^2(D)$. The choice of the

direction $v_{i,n}$ in (17) and (18) is the Polak-Ribière conjugate gradient direction [*Brodlie*, 1977], assuming the contrast does not change. Continue this algorithm until we again achieve $\varepsilon_{n-1} < \varepsilon$. The resulting values are taken as $u_i^{\text{initial}}$ and $\chi^{\text{initial}}$. At each step the constants $\alpha_n$ and $\beta_n$ are determined as described above. In our computations we choose $\varepsilon = 0.01$; however, this is arbitrary, and other choices could be made. No attempt was made to find the optimal $\varepsilon$, one that minimizes the number of iterations.

With these initial choices we run the algorithm of (9) with $v_{i,n}$ as in (17) and (18), and $d_n$ is taken in one of the two ways: if $\varepsilon_{n-1} \geq \varepsilon$, then $d_n$ is taken to be the gradient direction assuming that the fields do not change; that is,

$$g_n^d = -w_D \sum_{i=1}^{I} \bar{u}_{i,n-1} \bar{G}_D r_{i,n-1} + w_S \sum_{i=1}^{I} \bar{u}_{i,n-1} \bar{G}_S \rho_{i,n-1}, \tag{19}$$

whereas if $\varepsilon_{n-1} < \varepsilon$, we use the Polak-Ribière conjugate gradient direction [*Brodlie*, 1977]

$$d_n = g_n^d + \gamma_n^d d_{n-1}, \qquad \gamma_n^d = \frac{\langle g_n^d, g_n^d - g_{n-1}^d \rangle_D}{\|g_{n-1}^d\|_D^2}. \tag{20}$$

Continue the iteration until either $F_n$ meets a preset error criterion or ceases to change. In all examples considered we were able to drive the normalized error $F_n$ below 1.5% before it ceased changing with further iterations.

It should be pointed out that the choice of the update directions described here is geared to achieving reconstructions for high contrasts. When the contrast is low, not only is it unnecessary to use the sophisticated update directions for the field given by (17) and (18), it is actually beneficial to use the simpler update direction of the successive overrelaxation method for solving the direct problem. The update directions for the contrasts, however, should still be chosen as in (19) and (20). An explanation for this behavior is that for low contrasts the first few iterations in the successive overrelaxation method for solving the field equation converges faster than the corresponding number of iterations in the conjugate gradient method. Thus for low contrasts, corrections in the field based on the overrelaxation method are preferable to those based on the conjugate gradient methods. What constitutes a "low" contrast is determined by the rapidity of convergence of the overrelaxation method for solving the direct problem. In the inverse problem, such information may not be available, in which case, tests may have to be run using both choices for the field updates.

## NUMERICAL EXAMPLES

In actual numerical examples a discrete form of the algorithm was used. In these examples it was assumed that the unknown scatterer was located entirely within a test square of known dimension, although knowledge of the precise location within the test square was not assumed. This test square was partitioned into $J^2$ equal-sized subsquares, and the integrals over the domain $D$ in the algorithm were all carried over this test square. The position of the actual scatterer is determined as the support (nonzero values) of the reconstructed contrast. The domain integrals were approximated by assuming that the contrast and fields were constant on subsquares. The resulting integrals over subsquares were approximated by integrals over circles of equal area, which were calculated analytically [*Richmond*, 1965]. The discrete spatial convolutions of the $G_D$ operators were computed using fast Fourier transform routines [*Van den Berg*, 1984].

The measurement surface $S$ was chosen to be a circle containing the test domain. The incident fields were chosen to be line sources parallel to the axis of the scatterer considered as a cylinder in $\mathcal{R}^3$. These sources were taken to be equally spaced on the measurement circle, and the source locations were also chosen as discretization points on the circle. All integrals on $S$ were approximated by point collocation at the discretization points, that is, the rectangular rule with the integrand evaluated at the end point. The measured data were simulated by solving the direct scattering problem with a conjugate gradient method [*Van den Berg*, 1984]. The forward solver was run until a residual error criteria of $10^{-10}$ was met: that is, $n$ was taken large enough so that $w_D \sum_{i=1}^{I} \|r_{i,n}\| < 10^{-10}$. This inversion method is illustrated in a number of examples.

In the first two examples the test square was $d = 3\lambda$ on a side and was divided into $29 \times 29$ subsquares ($J = 29$). The measurement surface was a circle of radius $3\lambda$. There were 30 measurement stations equally spaced on the circle, each of which served in turn as the location of a line source ($I = 30$).

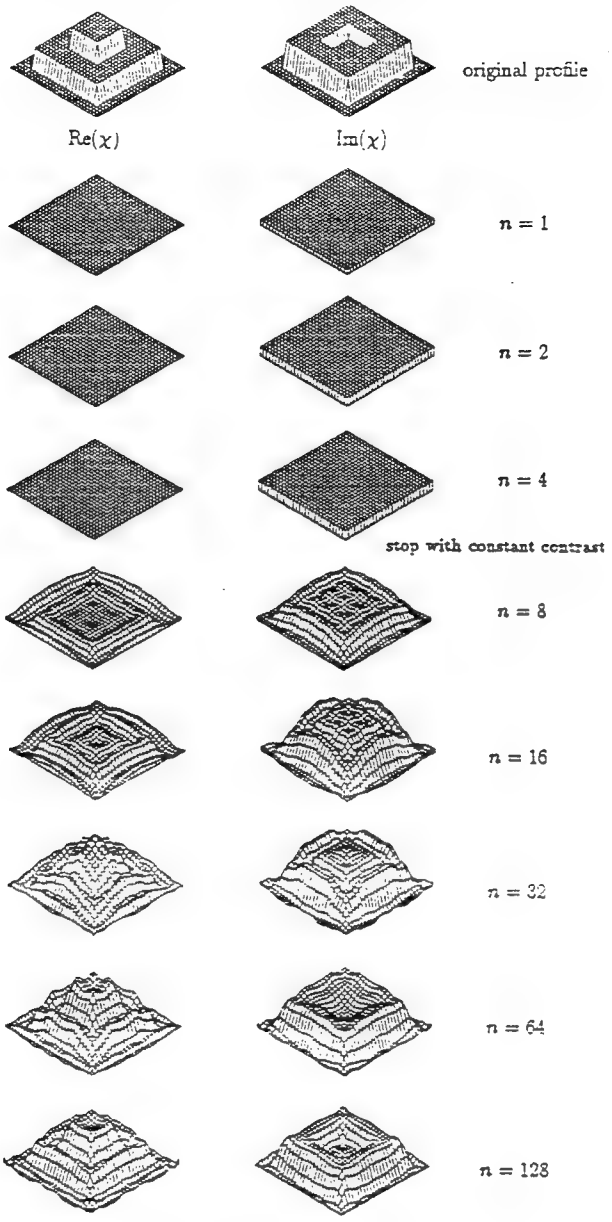



Fig. 1. Reconstruction of complex contrast.

In the first example the actual profile was inhomogeneous and complex, consisting of a square of dimension $\lambda \times \lambda$ with contrast $\chi = 0.6 + 2i$, surrounded by a larger square, $2\lambda \times 2\lambda$ with contrast $\chi = 0.3 + 0.4i$. Outside of this square the contrast was zero, so that the scattering object was smaller than the test square. The actual profile is shown in Figure 1. This example was also treated by *Kleinman and Van den Berg* [1992], who used the earlier approach. Here we use the algorithm described in the present paper. A constant $\chi^{\text{initial}}$ was first found using $d_n = 1$ until $\varepsilon_n < 0.01$. Then, the Polak-Ribière directions (19) and (20) were employed. The field update directions were always chosen to be those of the successive overrelaxation method, $v_{i,n} = r_{i,n-1}$, as the contrast was sufficiently low so that the Polak-Ribière directions (17) and (18) were not needed. The results of the recon-

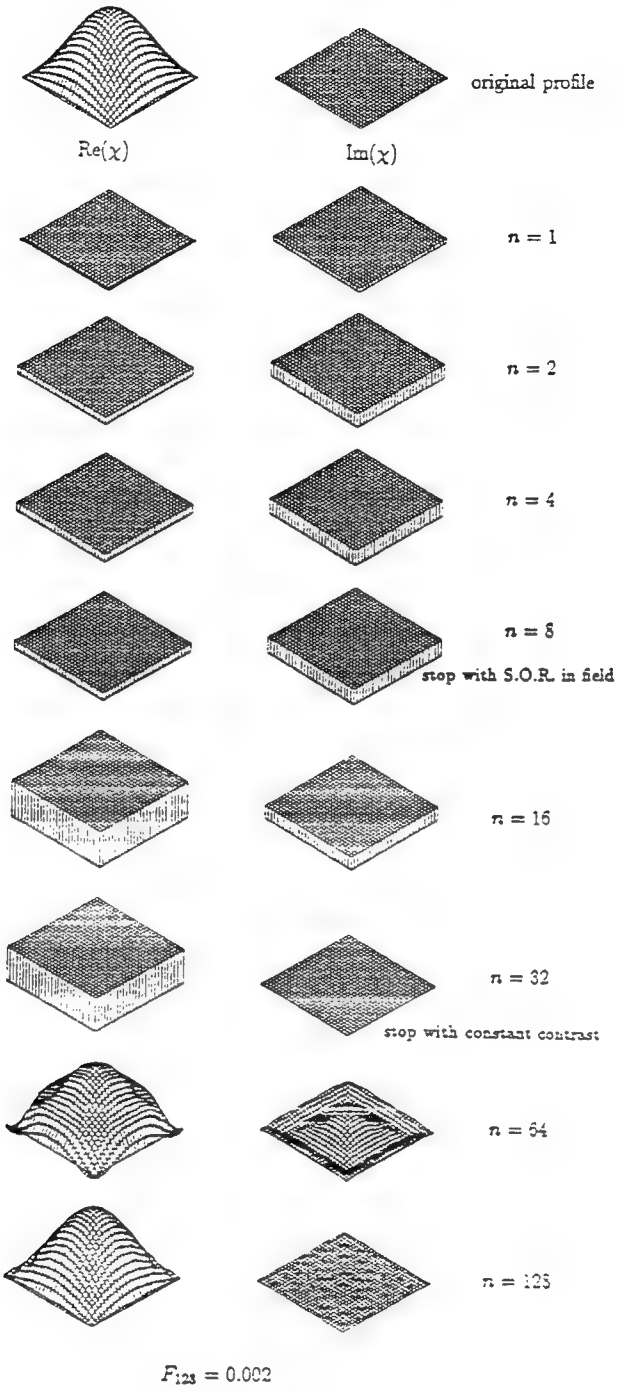



Fig. 2. Reconstruction of real contrast: limiting case of $kd = 6\pi$ and $\chi_{max} = 1$; 30 stations.

struction are shown in Figure 1, where the switching point is indicated. After 128 iterations, when $F_{128} = 0.004$, a reconstructed profile was obtained which required about 512 iterations using the earlier scheme. Thus the new definition of the update directions for the contrast resulted in a savings of approximately a factor of 4 in iterations.

In our second example we considered a higher

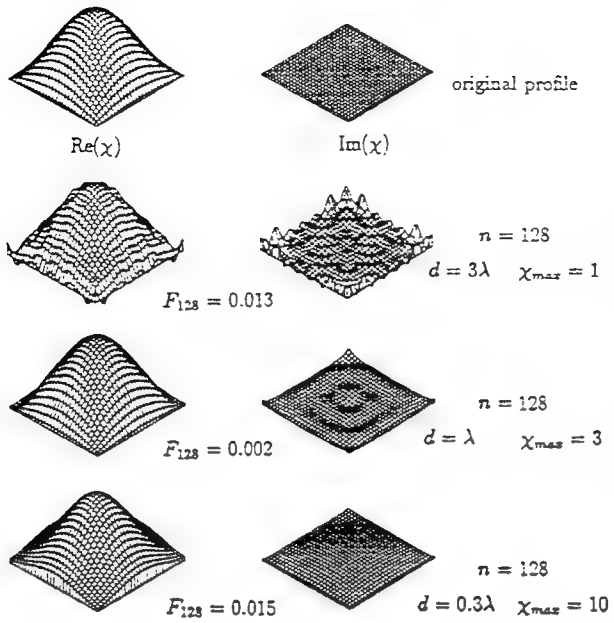



Fig. 3. Reconstruction of real contrast: limiting cases of $kd\chi_{max} = 6\pi$; 20 stations.

maximum contrast so that the Polak-Ribière directions were used in both field and contrast updates. The actual profile was sinusoidal in both $x$ and $y$, $\chi = \sin(\pi x/d) \sin(\pi y/d)$ for $0 < x, y < d = 3\lambda$, so that $kd|\chi_{max}| = 6\pi$. The actual profile and the reconstruction are shown in Figure 2. We start with the simplest scheme in which we use the field update directions of the successive overrelaxation method (SOR in Figure 2) $v_{i,n} = r_{i,n-1}$, and the constant contrast update directions $d_n = 1$. As shown in Figure 2, this simplest scheme was used until $n = 8$. Then with a switching criterion of $\varepsilon = 0.01$ the Polak-Ribière directions were used only for the field updates until $n = 32$ after which they were used to update both field and contrast. The original contrast was well constructed after 128 iterations at which point the cost functional had a value of $F_{128} = 0.002$. Experiments with higher values of $kd|\chi_{max}|$ indicated that the algorithm failed to reliably reconstruct the profile.

Additional examples were investigated to determine how object size and $|\chi_{max}|$ individually influence the reconstruction, as the amount of data diminishes. In these examples the test square was still divided into $29 \times 29$ subsquares, the profile was still sinusoidal, with $\chi = \chi_{max} \sin(\pi x/d) \sin(\pi y/d)$, and $kd|\chi_{max}| = 6\pi$. However, three different pairs of values of test square dimension $d$ and $\chi_{max}$ were tested: (1) $d = 3\lambda$, $\chi_{max} = 1$, (2) $d = \lambda$, $\chi_{max} = 3$, and (3) $d = 0.3\lambda$, $\chi_{max} = 10$. Moreover, the number of source and receiver stations was reduced to 20, $I = 20$. For each case the full method was employed using a switching criterion of $\varepsilon = 0.01$. The results after 128 iterations are shown in Figure 3. They show that the loss of data causes most instability at the shorter wavelengths, whereas spatial resolution diminishes at longer wavelengths. In obtaining the results in Figure 2 with 30 stations and those of Figure 3 for $d = 3\lambda$, $\chi_{max} = 1$, it was found that 56 iterations were needed to obtain the initial guess for the case with 20 stations, while 38 were needed with 30 stations. Moreover, the values of the functional in (11), which was to be minimized were $F_{128} = 0.013$ for $I = 20$ and $F_{128} = 0.002$ for $I = 30$, almost a factor of 10 smaller.

## CONCLUSIONS

An iterative method for complex profile reconstruction has been considerably refined to achieve significantly greater efficiency. These refinements have been described, and the limits of the new algorithm have been tested. The method combines the features of successive overrelaxation, gradient, and conjugate gradient methods to minimize a functional consisting of normalized errors in satisfying the field equation and the error in matching the measured data. The field equation serves as the regularizer for the ill-posed problem finding a function in $L_2(D)$ to minimize the error in solving (4). The nonlinear optimization problem is not linearized; however, the two components of the functional in (7) are treated somewhat separately. The algorithm was constructed to delay large changes in the contrast until the field was somewhat stable. This was the motivation for the separate treatment of the initial guesses as well as the subsequent switching in the algorithm based on the magnitude of the change in consecutive approximations of the field, $\varepsilon_n$. The numerical results presented here, as well as additional experiments indicate that the algorithm successfully reconstructs complex contrasts for $kd|\chi_{max}| \leq 6\pi$. To achieve reconstructions for large values of $\chi_{max}$, low-frequency measurements will not suffice to give reasonable resolution. Further work is directed toward extending the method to include measurements at more than one frequency to accommodate larger contrasts. The

algorithm appears to be stable with respect to noise. We checked the effect of introducing random noise equal to 10% of the maximum value of the data. The algorithm was run, and at each iteration the real and imaginary parts of the contrast were set equal to zero if negative values were obtained. The reconstructed profiles displayed a noisy distortion roughly equivalent to the magnitude of the noise.

*Acknowledgments.* This work was supported under NSF grant DMS-8912593, AFOSR grant 91-0277, NATO grant 0230/88, and a research grant from Schlumberger-Doll Research, Ridgefield, Connecticut.

## REFERENCES

Brodlie, K. W., Unconstrained minimization, in *The State of the Art in Numerical Analysis*,edited by D. A. H. Jacobs, chap. III, pp. 255–257, Academic, San Diego, Calif., 1977.

Chew, W. C., and Y. M. Wang, Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method, *IEEE Trans. Med. Imag., MI-9*, 218–225, 1990.

Colton, D., and P. Monk, The numerical solution of an inverse scattering problem for acoustic waves, *J. Appl. Math., 49*, 163–184, 1992.

Habashy, T. M., M. L. Oristaglio, and A. T. de Hoop, Simultaneous inversion of permittivity and conductivity employing a nonperturbative approach, in *SPIE 1767, Inverse Problems in Scattering and Imaging*, pp. 193–205, Society of Photo-Optical Instrumentation Engineers, Bellingham, Wash., 1992.

Isakov, V., On uniqueness in the inverse transmission scattering problem, *Commun. Partial Differential Equations, 15*, 1565–1587, 1990.

Joachimowicz, N., C. Pichot, and J. P. Hugonin, Inverse scattering: An iterative numerical method for electromagnetic imaging, *IEEE Trans. Antennas Propag., AP-39*, 1742–1752, 1991.

Kleinman, R. E., and P. M. van den Berg, Iterative methods for solving integral equations, in *Application of Conjugate Gradient Methods to Electromagnetics and Signal Analysis*, edited by T. K. Sarker, chap. 3, pp. 67–102, Elsevier Science, New York, 1991.

Kleinman, R. E., and P. M. van den Berg, A modified gradient method for two-dimensional problems in tomography, *J. Comp. Appl. Math., 42*, 17–35, 1992.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes, The Art of Scientific Computing*, Cambridge University Press, New York, 1986.

Richmond, J. H., Scattering by a dielectric cylinder of arbitrary cross section shape, *IEEE Trans. Antennas Propag., AP-13*, 334–341, 1965.

Roger, A., Newton-Kantorovitch algorithm applied to an electromagnetic inverse scattering problem, *IEEE Trans. Antennas Propag., AP-29*, 232–238, 1981.

Van den Berg, P. M., Iterative computational techniques in scattering based upon the integrated square error criterion, *IEEE Trans. Antennas Propag., AP-32*, 1063–1071, 1984.

R. E. Kleinman, Center for the Mathematics of Waves, Department of Mathematical Sciences, University of Delaware, Newark, DE 19716.

P. M. van den Berg, Laboratory of Electromagnetic Research, Faculty of Electrical Engineering, Delft University of Technology, PO Box 5031, 2600 GA Delft, The Netherlands.

*Radio Science*, Volume 29, Number 4, Pages 1157–1169, July–August 1994

# Two-dimensional location and shape reconstruction

R. E. Kleinman

Center for the Mathematics of Waves, Department of Mathematical Sciences, University of Delaware, Newark

P. M. van den Berg

Laboratory of Electromagnetic Research, Faculty of Electrical Engineering, Delft University of Technology, Delft, Netherlands

**Abstract.** A method for reconstructing the location and the shape of a bounded impenetrable object from measured scattered field data is presented. The algorithm is, in principle, the same as that used for reconstructing the conductivity of a penetrable object and uses the fact that for high conductivity the skin depth of the scatterer is small, in which case the only meaningful information produced by the algorithm is the boundary of the scatterer. A striking increase in efficiency is achieved by incorporating into the algorithm the fact that for large conductivity the contrast is dominated by a large positive imaginary part. This fact, together with the knowledge that the scatterer is constrained in some test domain, constitute the only a priori information about the scatterer that is used. There are no other implicit assumptions about the location, connectivity, convexity, or boundary conditions. Some refinements of the algorithm which reduce the number of points at which the unknown function is updated are incorporated to further increase efficiency. Results of a number of numerical examples are presented which demonstrate the effectiveness of the location and shape reconstruction algorithm.

## Introduction

Among the many inverse problems of current interest there are two general classes of primary concern in acoustics: electromagnetics and seismics. One class involves the determination of the constitutive parameters of a penetrable scatterer (e.g., local sound speed, index of refraction, conductivity), while the second class is concerned with determining the shape of the boundary of an impenetrable scatterer. In both cases the location and orientation of the scatterer is also of interest. The data from which these reconstructions are attempted consist of a knowledge of how the object perturbs known exciting fields at points exterior to the object. When the exciting field is one or more incident waves it is standard to use a knowledge of whether the object is penetrable or impenetrable as a priori information in designing reconstruction algorithms. Even when the methods stem from the same mathematical approach such as the use of complete families [*Angell et al.*, 1986, 1989] or Herglotz wave functions [*Colton and Monk*, 1987; *Colton and Kress*, 1992], the algorithms specifically incorporate information as to whether the scattering object is penetrable or not.

The present paper describes a method for reconstructing the location and shape of the boundary of an impenetrable object without making the a priori assumption of impenetrability. In fact, the algorithm is precisely the same as that used for reconstructing the conductivity of a penetrable object and uses the fact that for high conductivity the skin depth of the scatterer is small, in which case the only meaningful information produced by the algorithm is the boundary of the scatterer.

This work is a further development of the method described by *Kleinman and Van den Berg* [1992] for reconstructing the complex index of refraction of an unknown scatterer from a knowledge of the field scattered when the object is illuminated successively by a number of different excitations. The method consists of casting the inverse problem as an optimization problem in which the cost functional is the sum of two terms: One is the defect in matching measured (actual or synthetic) field data



Paper number 93RS03445.
0048-6604/94/93RS-03445$08.00

with the field scattered by a body with a particular index of refraction, and the second is the error in satisfying the equations of state, integral equations for the field produced in the body by each excitation. The index and the fields are each updated by a linear iterative method in which the updating directions are weighted by parameters which are determined by minimizing the cost functional. A simple choice for the updating directions was made by *Kleinman and Van den Berg* [1992] which sufficed to enable some remarkable reconstructions. A more sophisticated choice of updating directions was described by *Kleinman and Van den Berg* [1993] which resulted in a more efficient algorithm. This algorithm was tested to determine its limits, and a rough estimate of the upper limit of reconstructibility was found to be $kd|\chi_{\max}| = 6\pi$, where $k$ is the wavenumber, $d$ is the diameter of the domain of investigation, and $|\chi_{\max}|$ is the maximum modulus of the contrast $\chi$ defined in terms of the index of refraction to be $\chi = n^2 - 1$. This limit was determined from examples of contrasts with nonzero real part and is definitely dependent on the algorithm, as is clear since in this paper we show that by changing the algorithm, considerable higher contrasts are reconstructed.

Consider the scattering object to be an inhomogeneous lossy dielectric cylinder with relative permeability equal to one and of arbitrary cross section imbedded in free space. When the incident excitation consists of electromagnetic waves with the electric vector **E** polarized along the cylinder axis, then the contrast is given by $\chi = \varepsilon_r - 1 - i\sigma/\omega\varepsilon_0$, where $\varepsilon_r$ is the relative permittivity of the object, $\sigma$ is the conductivity, $\varepsilon_0$ is the free space permittivity, and $\omega$ is the angular frequency (time factor is $\exp(-i\omega t)$). No attempt was made to incorporate the information that $\varepsilon_r$ and $\sigma/\omega\varepsilon_0$ are nonnegative quantities into the previously described algorithm. Recently, *Habashy et al.* [1992, 1994] demonstrated that by explicitly incorporating this a priori information into a different algorithm, contrasts considerably higher than $kd|\chi_{\max}| = 6\pi$ could be reconstructed. In the present paper we combine this idea of enforcing positivity together with the physically motivated approximation that for large $\sigma/\omega\varepsilon_0$, $\chi \sim i\sigma/\omega\varepsilon_0$ even if $\varepsilon_r \neq 1$, to modify our algorithm appropriately. This new algorithm is tested using synthetic data and is shown to be extremely effective in reconstructing the location and the boundary of the scatterer. The details of the method are given by *Kleinman and Van den Berg* [1992], and here we will present only the essential steps and include the changes needed to enforce the a priori positivity constraint.

## Description of the Two-Dimensional Problem

Assume that a two-dimensional conducting obstacle $D$ is irradiated successively by a number of known incident fields $u_j^{\text{inc}}$, $j = 1, \cdots, J$. For each excitation the direct scattering problem may be reformulated as the domain integral equation

$$L_{(\chi)}u_j(\mathbf{p}) := u_j(\mathbf{p}) - G_D\chi u_j(\mathbf{p}) = u_j^{\text{inc}}, \quad \mathbf{p} \in D, \tag{1}$$

where

$$G_D\chi u_j(\mathbf{p}) := k^2 \int_D G(\mathbf{p}, \mathbf{q})\chi(\mathbf{q})u_j(\mathbf{q})\, dv_{\mathbf{q}}, \quad \mathbf{p} \in D, \tag{2}$$

and

$$G(\mathbf{p}, \mathbf{q}) = \frac{i}{4} H_0^{(1)}(k|\mathbf{p} - \mathbf{q}|). \tag{3}$$

Here $u_j$ is the total field corresponding to the incident field $u_j^{\text{inc}}$, $k$ is the wavenumber, $\chi$ is taken to be equal to $i\zeta^2$ for real $\zeta$ ($\zeta^2 = \sigma/\omega\varepsilon_0$), and $\mathbf{p}$ and $\mathbf{q}$ are position vectors. $G(\mathbf{p}, \mathbf{q})$ is the free space Green's function in two dimensions. $G_D$ is an operator mapping $L^2(D)$ (square integrable functions in $D$) into itself. If $S$ is a surface enclosing $D$, then the scattered field $u_j^{\text{sct}}$ on $S$ is given by $G_S\chi u_j$ where $G_S$ is the same operator defined in (2), except the field point $\mathbf{p}$ now lies on $S$. Hence $G_S$ is an operator mapping $L^2(D)$ into $L^2(S)$. We assume that $u_j^{\text{sct}}$ is measured on $S$ and denoted by $f_j(\mathbf{p})$, $\mathbf{p} \in S$, the measured data for each excitation $j$, $j = 1, \cdots, J$. The conductivity reconstruction problem is that of finding $\chi$ for given $f_j$ or solving the equations

$$G_S\chi u_j(\mathbf{p}) = f_j(\mathbf{p}), \quad \mathbf{p} \in S, \quad j = 1, \cdots, J, \tag{4}$$

for $\chi$, subject to the additional condition that $u_j$ and $\chi = i\zeta^2$ satisfy (1) in $D$ for each $j$. Thus there are two error measurements involved; the first is the defect in matching the measured data in $L_2(S)$, namely,

$$F_{S,j} = \|f_j - iG_S\zeta^2 u_j\|_S^2 \tag{5}$$

and the second is the error in the state equation in $L_2(D)$

$$F_{D,j} = \|u_j^{\text{inc}} - L_{(i\zeta^2)}u_j\|_D^2, \tag{6}$$

where the subscripts $S$ and $D$ are included in the norm $\|\cdot\|$ and later the inner product $\langle \cdot, \cdot \rangle$ in $L^2$ to indicate the domain of integration. Rather than seek $\zeta$ to minimize the data error subject to the state error as a constraint, we combine these two error measures into one normalized cost functional

$$F = w_D \sum_{j=1}^{J} \|u_j^{\text{inc}} - L_{(i\zeta^2)}u_j\|_D^2 + w_S \sum_{j=1}^{J} \|f_j - iG_S\zeta^2 u_j\|_S^2, \tag{7}$$

where

$$w_D = \left(\sum_{j=1}^{J} \|u_j^{\text{inc}}\|_D^2\right)^{-1} \quad \text{and} \quad w_S = \left(\sum_{j=1}^{J} \|f_j\|_S^2\right)^{-1}. \tag{8}$$

In practice we approximate $\zeta$ and $u_j$ in subspaces of $L^2(D)$, and hence the guiding principle is to simultaneously seek $\zeta$ and $u_j$ in a way which minimizes $F$. The functional $F_{S,j}$ is used as the starting point for many inversion algorithms. Combining it with $F_{D,j}$ as in (7) is less common but has been used before [see *Kleinman and Van den Berg*, 1992, and references therein]. Recently, *Sabbagh and Lautzenheiser* [1993] used the same functional in an inversion algorithm. Domain functionals have also been used in impedance tomography [e.g., *Wexler et al.*, 1985].

## Inversion Algorithm

The basic idea underlying the inversion algorithm is to combine a gradient type of algorithm to iteratively solve the direct scattering problem together with a similar algorithm for solving the ill-posed inverse problem. *Sabbagh and Lautzenheiser* [1993] attempted to do this by constructing one unknown vector consisting of all the fields and the unknown contrast and updating with one gradient direction and one coefficient. Our approach differs in that we update each field and the contrast separately. Specifically, we propose the iterative construction of sequences $\{u_{j,n}\}$ and $\{\zeta_n\}$ as follows:

$$u_{j,n} = u_{j,n-1} + \alpha_n v_{j,n}, \qquad \zeta_n = \zeta_{n-1} + \beta_n \xi_n, \tag{9}$$

$$n = 1, 2, \cdots.$$

For each $n$, the functions $v_{j,n}$ and $\xi_n$ are update directions for the functions $u_{j,n}$ and $\zeta_n$, respectively, while the complex parameter $\alpha_n$ and the real parameter $\beta_n$ are weights to be determined. The residual errors at each step in the state equation and data equation are defined as

$$r_{j,n} = u_j^{\text{inc}} - L_{(i\zeta_n^2)}u_{j,n}, \qquad \rho_{j,n} = f_j - iG_S\zeta_n^2 u_{j,n}, \tag{10}$$

and the value of the cost functional at the $n$th step is

$$F_n = w_D \sum_{j=1}^{J} \|r_{j,n}\|_D^2 + w_S \sum_{j=1}^{J} \|\rho_{j,n}\|_S^2, \tag{11}$$

where the residuals satisfy the recursive relations

$$\begin{aligned} r_{j,n} = r_{j,n-1} &- \alpha_n L_{(i\zeta_{n-1}^2)}v_{j,n} + 2i\beta_n G_D \zeta_{n-1}\xi_n u_{j,n-1} \\ &+ 2i\alpha_n\beta_n G_D \zeta_{n-1}\xi_n v_{j,n} + i\beta_n^2 G_D \xi_n^2 u_{j,n-1} \\ &+ i\alpha_n\beta_n^2 G_D \xi_n^2 v_{j,n}, \end{aligned} \tag{12}$$

$$\begin{aligned} \rho_{j,n} = \rho_{j,n-1} &- i\alpha_n G_S \zeta_{n-1}^2 v_{j,n} - 2i\beta_n G_S \zeta_{n-1}\xi_n u_{j,n-1} \\ &- 2i\alpha_n\beta_n G_S \zeta_{n-1}\xi_n v_{j,n} - i\beta_n^2 G_S \xi_n^2 u_{j,n-1} \\ &- i\alpha_n\beta_n^2 G_S \xi_n^2 v_{j,n}. \end{aligned} \tag{13}$$

Substitution of these expressions in (11) results in an expression involving terms determined at the $(n-1)$st step, the directions $\xi_n$ and $v_{j,n}$, and the two parameters $\alpha_n$ and $\beta_n$. Once the directions $\xi_n$ and $v_{j,n}$ are chosen we have a nonlinear expression in $\alpha_n$ and $\beta_n$. As *Kleinman and Van den Berg* [1992] did, the values of the parameters $\alpha_n$ and $\beta_n$ are determined by requiring $F_n$ to be a minimum. This minimization of the quantity $F_n$ is accomplished using the Fletcher-Reeves-Polak-Ribière conjugate gradient method [*Press et al.*, 1986]. The starting value for $\alpha_n$ is obtained by taking $\beta_n = 0$ and minimizing $F_n$, while the starting value for $\beta_n$ is found by setting $\alpha_n = 0$, neglecting the terms of $\beta_n^2$ and again minimizing $F_n$. The essential ingredients remaining are the update directions $v_{j,n}$ and $\xi_n$ and the initial choices $u_{j,0}$ and $\zeta_0$.

## Update Directions

As the update direction for the field we take the direction

$$v_{j,n} = g^v_{j,n} + \gamma^v_n v_{j,n-1}, \quad \gamma^v_n = \frac{\sum_{j=1}^{J} \langle g^v_{j,n}, g^v_{j,n} - g^v_{j,n-1} \rangle_D}{\sum_{j=1}^{J} \|g^v_{j,n-1}\|^2_D}. \tag{14}$$

Here $g^v_{j,n}$ is the gradient of $F$ (see equation (7)), with respect to changes in the field $u_j$, evaluated at the $(n-1)$st step, that is,

$$g^v_{j,n} = w_D(r_{j,n-1} + i\zeta^2_{n-1}\bar{G}_D r_{j,n-1}) - iw_S\zeta^2_{n-1}\bar{G}_S\rho_{j,n-1}, \tag{15}$$

where the overbar denotes complex conjugate,

$$\bar{G}_D r_{j,n-1}(\mathbf{q}) = k^2 \int_D \bar{G}(\mathbf{p}, \mathbf{q}) r_{j,n-1}(\mathbf{p})\, dv_{\mathbf{p}}, \quad \mathbf{q} \in D, \tag{16}$$

and

$$\bar{G}_S\rho_{j,n-1}(\mathbf{q}) = k^2 \int_S \bar{G}(\mathbf{p}, \mathbf{q})\rho_{j,n-1}(\mathbf{p})\, dv_{\mathbf{p}}, \quad \mathbf{q} \in D. \tag{17}$$

$\bar{G}_D$ is an operator mapping $L^2(D)$ into itself, while $\bar{G}_S$ is a map from $L^2(S)$ to $L^2(D)$. The choice of the direction $v_{j,n}$ in (14)–(15) is the Polak-Ribière conjugate gradient direction [*Brodlie*, 1977] assuming the contrast does not change.

As the update direction for the contrast we take the direction

$$\xi_n = g^\xi_n + \gamma^\xi_n \xi_{n-1}, \qquad \gamma^\xi_n = \frac{\langle g^\xi_n, g^\xi_n - g^\xi_{n-1} \rangle_D}{\|g^\xi_{n-1}\|^2_D}, \tag{18}$$

where $g^\xi_n$ is the gradient of $F$ with respect to changes in $\xi$, evaluated at the $(n-1)$st step, that is,

$$g^\xi_n = 2\zeta_{n-1}\,\mathrm{Im}\left[ w_D \sum_{j=1}^{J} \bar{u}_{j,n-1}\bar{G}_D r_{j,n-1} - w_S \sum_{j=1}^{J} \bar{u}_{j,n-1}\bar{G}_S\rho_{j,n-1} \right]. \tag{19}$$

The choice of the direction $\xi_n$ in (18)–(19) is the Polak-Ribière conjugate gradient direction [*Brodlie*, 1977] assuming the fields do not change. Note that the contrast gradient $g^\xi_n(\mathbf{q})$ vanishes for zero values of $\zeta_{n-1}(\mathbf{q})$. We therefore cannot start the iterative scheme with a zero estimate for $\zeta_0$, and a more careful choice must be made.

## Initial Choice

The initial choice is determined from a guess of the contrast sources

$$w_j(\mathbf{q}) = i\zeta^2(\mathbf{q})u_j(\mathbf{q}), \quad \mathbf{q} \in D, \tag{20}$$

that follow from the linear data equation

$$G_S w_j(\mathbf{p}) = f_j(\mathbf{p}), \quad \mathbf{p} \in S. \tag{21}$$

In contrast to *Habashy et al.* [1992], we do not solve this first-kind integral equation, but we take an estimate

$$w_{j,0}(\mathbf{q}) = \gamma\bar{G}_S f_j(\mathbf{q}). \tag{22}$$

The constant $\gamma$ is determined by minimizing the error, see (5),

$$\sum_{j=1}^{J} F_{S,j} = \sum_{j=1}^{J} \|f_j - G_S w_{j,0}\|^2_S = \sum_{j=1}^{J} \|f_j - \gamma G_S\bar{G}_S f_j\|^2_S. \tag{23}$$

This leads to

$$\gamma = \frac{\sum_{j=1}^{J} \langle f_j, G_S\bar{G}_S f_j \rangle_S}{\sum_{j=1}^{J} \|G_S\bar{G}_S f_j\|^2_S}. \tag{24}$$

With the initial estimate for the contrast sources $w_{j,0}$, an initial estimate for the fields $u_{j,0}$ follows from the state equation (2) as

$$u_{j,0}(\mathbf{p}) = u^{\mathrm{inc}}_j(\mathbf{p}) + G_D w_{j,0}(\mathbf{p}), \quad \mathbf{p} \in D. \tag{25}$$

Once the initial estimates for the contrast sources and the fields have been determined an initial estimate for the contrast $\zeta_0$ ($\zeta_0 > 0$) follows from a minimization procedure of the error in the constitutive relationship (20) [see *Habashy et al.*, 1994]. For the initial estimates this relation is rewritten as

$$\mathrm{Im}\,[w_{j,0}(\mathbf{q})\bar{u}_{j,0}(\mathbf{q})] = \zeta^2_0(\mathbf{q})|u_{j,0}(\mathbf{q})|^2. \tag{26}$$

In order to meet this relation for all $j$ we use the ideas of *Kohn and McKenney* [1990] and minimize the cost function

$$\mathcal{F}(\mathbf{q}) = \sum_{j=1}^{J} \left( \frac{1}{\zeta_0(\mathbf{q})} \frac{\operatorname{Im}\left[w_{j,0}(\mathbf{q})\bar{u}_{j,0}(\mathbf{q})\right]}{|u_{j,0}(\mathbf{q})|} - \zeta_0(\mathbf{q})|u_{j,0}(\mathbf{q})| \right)^2$$

$$= \sum_{j=1}^{J} \left\{ \frac{1}{\zeta_0^2(\mathbf{q})} \frac{\{\operatorname{Im}\left[w_{j,0}(\mathbf{q})\bar{u}_{j,0}(\mathbf{q})\right]\}^2}{|u_{j,0}(\mathbf{q})|^2} + \zeta_0^2(\mathbf{q})|u_{j,0}(\mathbf{q})|^2 \right.$$

$$\left. - 2 \operatorname{Im}\left[w_{j,0}(\mathbf{q})\bar{u}_{j,0}(\mathbf{q})\right] \right\}. \tag{27}$$

Note that only the first two terms in the second expression depend on $\zeta_0$. Minimization of this expression yields

$$\zeta_0^2(\mathbf{q}) = \left( \frac{\sum_{j=1}^{J} \frac{\{\operatorname{Im}\left[w_{j,0}(\mathbf{q})\bar{u}_{j,0}(\mathbf{q})\right]\}^2}{|u_{j,0}(\mathbf{q})|^2}}{\sum_{j=1}^{J} |u_{j,0}(\mathbf{q})|^2} \right)^{1/2}, \quad \mathbf{q} \in D. \tag{28}$$

With the expressions of (25) and (28) the initial estimates for $u_{j,0}$ and $\zeta_0$ have been determined, and the iterative scheme is now completely defined.

## Numerical Examples

In actual numerical examples a discrete form of the algorithm was used. In these examples it was assumed that the unknown scatterer was located entirely within a test square of known dimension although knowledge of the precise location within the test square was not assumed. This test square was partitioned into equal-sized subsquares, and the integrals over the domain $D$ in the algorithm were all carried over this test square. The position of the actual scatterer is determined as the support (nonzero values) of the reconstructed contrast. The domain integrals were approximated by assuming that the contrast and fields were constant on subsquares. The resulting integrals over subsquares were approximated by integrals over circles of equal area which were calculated analytically [*Richmond*, 1965]. The discrete spatial convolutions of the $G_D$ operators were computed using fast Fourier transform routines [*Van den Berg*, 1984].

The measurement surface $S$ is chosen to be a circle containing the test domain. We assume that the radius of this circle is large enough so that the far-field approximation of (4) may be employed, and the far-field coefficient is the quantity of interest so that the dependence on the radius is removed. In that case the data may be written as

$$f_j(\mathbf{p}) \sim \frac{i}{4} \left( \frac{2}{\pi k|\mathbf{p}|} \right)^{1/2} \exp\left( ik|\mathbf{p}| - i\frac{\pi}{4} \right) f_j^{\infty}(\hat{\mathbf{p}}), \tag{29}$$

and the data equation (4) may be replaced by

$$\int_D \exp(-ik\hat{\mathbf{p}} \cdot \mathbf{q})\chi(\mathbf{q})u_j(\mathbf{q})\, dv_{\mathbf{q}} = f_j^{\infty}(\hat{\mathbf{p}}), \quad \hat{\mathbf{p}} \in S, \tag{30}$$

where $\hat{\mathbf{p}}$ is the unit vector in the direction of observation and $S$ now denotes the space of these unit vectors, the unit circle. Further, $f_j^{\infty}(\hat{\mathbf{p}})$ is the measured far-field data. In the examples we measure the far-field at 30 stations equally spaced around the object. Each of the stations serves in turn as the location of a source ($J = 30$), and the incident fields can be approximated as plane waves. All integrals on $S$ were approximated by point collocation at the discretization points, that is, the rectangular rule with the integrand evaluated at the end point. The measured data were simulated by solving the direct scattering problem for an impenetrable circular cylinder. The analytic solution in terms of Bessel functions has been employed. The radius $a$ of this circular cylinder was 0.015 m. Our reconstruction of the location and the shape of this circular cylinder is illustrated in a number of examples.

### Example 1

In the first example the test square was divided into 31 × 31 subsquares of 0.003 × 0.003 m$^2$. The wavelength is $\lambda = 0.090$ m, so that $ka = \pi/3$. The measured data were calculated for the cylinder with origin at the center of the test square. We then solved the inverse problem using the algorithm described in the previous sections, and the error $F_n^{1/2}$ is plotted in Figure 1 (solid line). Although it is very hard to minimize the error in the fields inside the impenetrable object, we are still able to reach an error less than a few percent. Some surface plots of the reconstructed profiles (the imaginary part of the contrast $\chi$) are presented in Figure 2. We indeed observe that after a relatively small number of iterations only the boundary of the object becomes

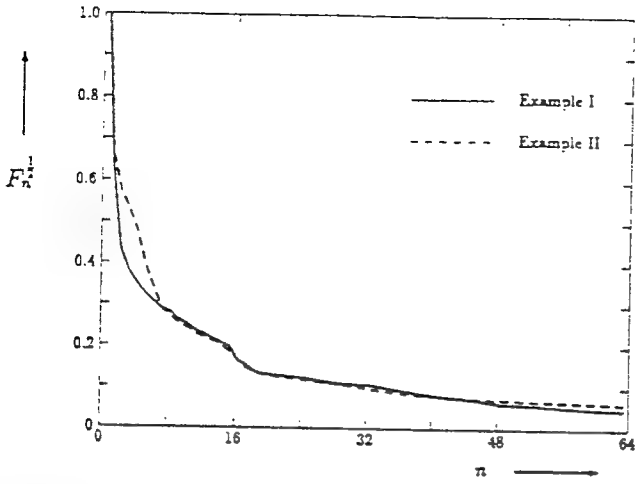



**Figure 1.** The error $F_n^{1/2}$ as a function of the number of iterations.

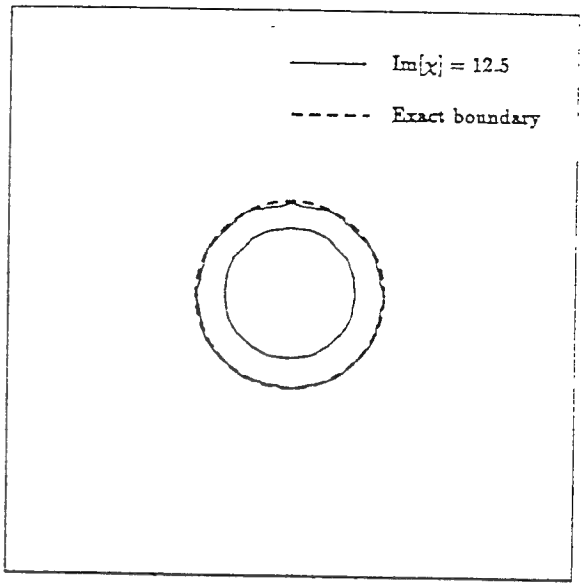



**Figure 3.** Comparison between the reconstructed boundary and the exact one of example 1 ($n = 64$).

visible, and the location and shape of the object can be estimated. Specifically, we observe that after ~16 iterations the imaginary part of the contrast at the boundary becomes larger than six, and only the contrast at the boundary of the object remains increasing when we increase the number of iterations. At 64 iterations the contrast at the boundary has reached values from 10 up to 25! This result has also been presented in Figure 3, where we have plotted the contour lines Im $[\chi] = 12.5$. The exact location of the boundary of the object is indicated by the dashed circle. The outer contour line almost coincides with the exact boundary, and we choose this as the reconstructed boundary.

### Example 2

In the second example we consider a smaller wavelength: $\lambda = 0.030$ m, so that $ka = \pi$. We then solved the inverse problem, and the error $F_n^{1/2}$ is plotted in Figure 1 (dashed line). Some surface plots of the reconstructed profiles (the imaginary part of

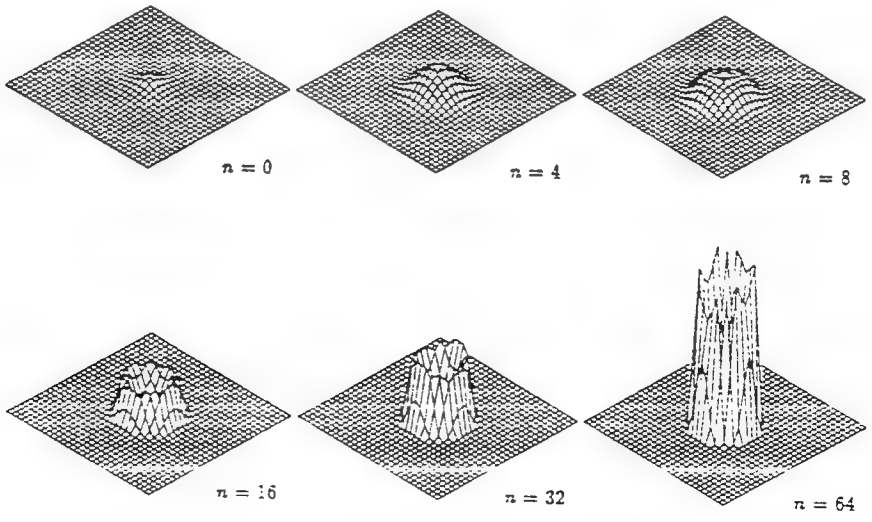



**Figure 2.** The reconstructed imaginary values of the contrast for example 1. At $n = 64$ the largest value is 25.3.

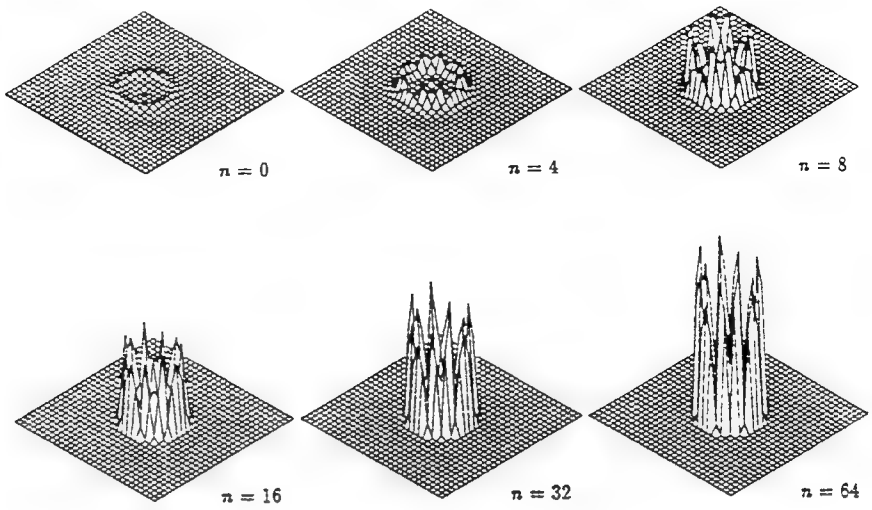



**Figure 4.** The reconstructed imaginary values of the contrast for example 2. At $n = 64$ the largest value is 6.02.

the contrast $\chi$) are presented in Figure 4. After only four iterations the boundary of the object is clearly visible. Specifically, we observe that after about eight iterations the imaginary part of the contrast at the boundary becomes larger than one and only the contrast at the boundary of the object remains increasing when we increase the number of iterations. After 64 iterations the contrast at the boundary has reached values from 2 up to 6. This result is also presented in Figure 5, where we have plotted the contour lines Im $[\chi] = 2.5$. The exact location of the boundary of the object is indicated by the dashed circle. The outer contour line approximates the boundary very well.

### Example 3

In the third example we still have $\lambda = 0.030$ m, however, the measured data were calculated for a cylinder located close to a corner of the test square. The reconstruction is shown in Figure 6. It shows that our scheme not only approximates the boundary of the object very well, but also the location is determined precisely. This is stressed in Figure 7, where after 64 iterations the contour lines Im $[\chi] = 2.5$ have been plotted. Again, the exact boundary of the object is indicated by the dashed circle.

## Bounded Contrast Reconstruction

In our examples we have seen that our scheme indeed reconstructs the location and the shape of an impenetrable object by reconstructing the imaginary contrast at the boundary. However, the reconstructed contrast at the boundary becomes highly oscillatory after a couple of iterations. The peaks appear to increase with the number of iterations, and it becomes difficult to choose the level value of the contour that estimates the boundary of the object. We therefore adopt a slightly modified reconstruction scheme. First of all we have observed

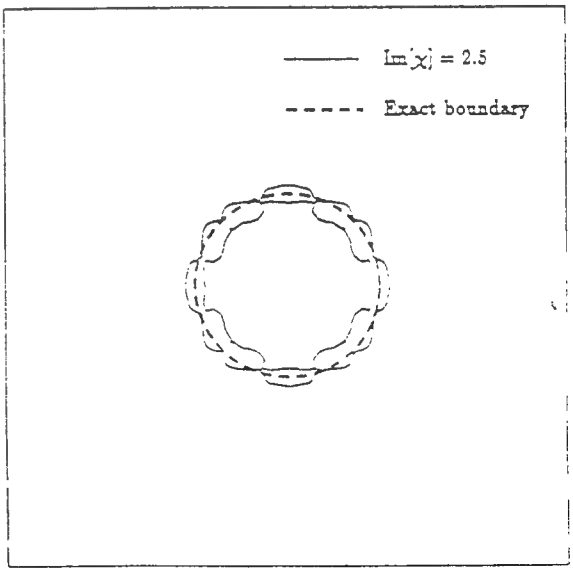



**Figure 5.** Comparison between the reconstructed boundary and the exact one of example 2 ($n = 64$).

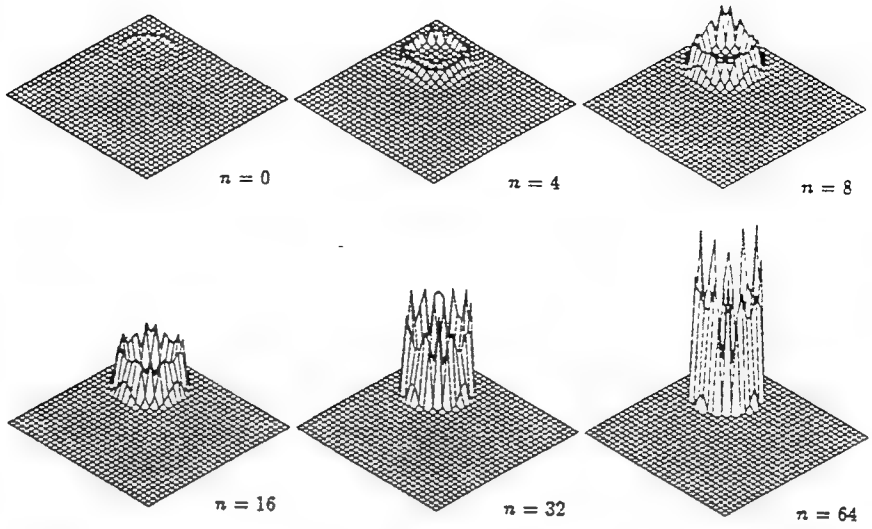



**Figure 6.** The reconstructed imaginary values of the contrast for example 3. At $n = 64$ the largest value is 8.87.

that there is no improvement in locating the boundary after the contrast has reached a value such that the penetration depth of the wavefield is of the order of the mesh width in the testing domain. The visualization of the boundary of the object is improved when we impose an upper bound to the reconstructed contrast in such a way that the penetration depth of the wavefield is not less than three times the mesh width. This factor is chosen to provide a reconstructed object such that we observe a "boundary wall" with a thickness of two or three times the mesh width. We therefore require that the interior (complex) wavenumber $k_i(\mathbf{q})$ satisfies the condition

$$\text{Im}\,[k_i(q)]3\Delta \leq 1, \tag{31}$$

where $\Delta$ is the side length of a subsquare of the test domain. From (31) and the fact that $\chi + 1 = k_i^2/k^2$, it follows that the maximum reconstructed contrast $\chi_{\max}$ follows from the relation

$$\text{Im}\,[(1 + \chi_{\max})^{1/2}] = \frac{1}{3k\Delta}. \tag{32}$$

The value of $\chi_{\max}$ is assumed to be pure imaginary and is determined numerically. If at some point in the iteration the reconstructed contrast is larger than $\chi_{\max}$, the contrast is replaced by $\chi_{\max}$. In view of this modification, the residual errors have to be recomputed and the iterative scheme restarts with new contrast directions $\xi_n$. By enforcing the contrast gradients to be zero in all the points $\mathbf{q}$, where the contrast is equal to $\chi_{\max}$, the contrast directions vanish in these points, and no updating of the contrast takes place in these points. Operating in this way, the scheme is able to "concentrate" on updating the contrast at the remaining points. This accelerates the reconstruction and visualization of

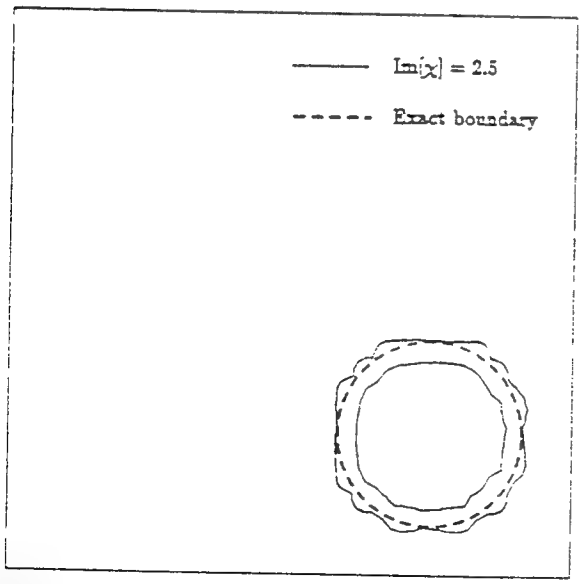



**Figure 7.** Comparison between the reconstructed boundary and the exact one of example 3 ($n = 64$).

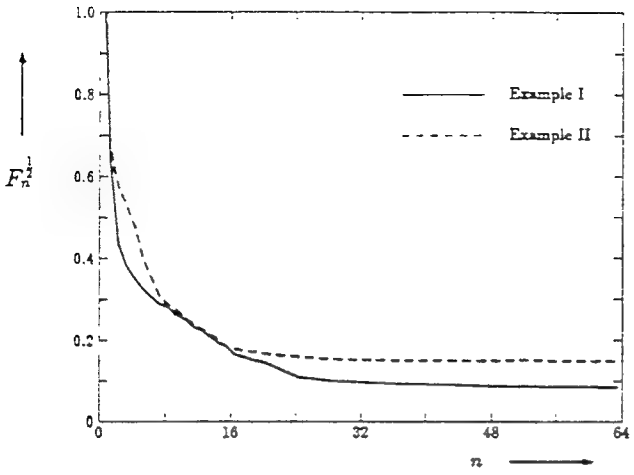



**Figure 8.** The error $F_n^{1/2}$ as a function of the number of iterations when the maximum reconstructed contrast is constrained.

the boundary of the object. We illustrate this procedure for our three examples.

### Example 1

In the first example with a wavelength of $\lambda = 0.090$ m and a side length of a subsquare of 0.003 m, condition (32) says that the maximum reliably reconstructed contrast amounts to $\chi_{max} = i5.97$. We then solved the inverse problem with this upper limit, and the error $F_n^{1/2}$ is plotted in Figure 8 (solid line). Some surface plots of the reconstructed profiles (the imaginary part of the contrast $\chi$) are presented in Figure 9. After 32 iterations we observe no substantial improvement in the reconstruction, as is seen by examining the reconstructed profile after 128 iterations. Comparing Figures 1 and 8, the error $F_{64}^{1/2}$ is now much larger, but this is mainly due to the mismatch in the fields inside the object. Relaxing our constraint on the maximum value of the contrast will decrease this error, but it does not yield better reconstruction of the boundary of the object. The reconstruction of the boundary is visualized in Figure 10, where we have plotted the contour lines $\chi = \chi_{max}$. The exact location of the boundary of the object is indicated by the dashed circle.

### Example 2

In the second example with a wavelength of $\lambda = 0.030$ m and a side length of a subsquare of 0.003 m, condition (32) says that the maximum reliably reconstructed contrast amounts to $\chi_{max} = i1.20$. We then solved the inverse problem with this upper limit, and the error $F_n^{1/2}$ is plotted in Figure 8 (dashed line). Some surface plots of the reconstructed profiles (the imaginary part of the contrast $\chi$) are presented in Figure 11. The result after 64 iterations is also presented in Figure 12, where we have plotted the contour lines $\chi = \chi_{max}$. The exact location of the boundary of the object is indicated

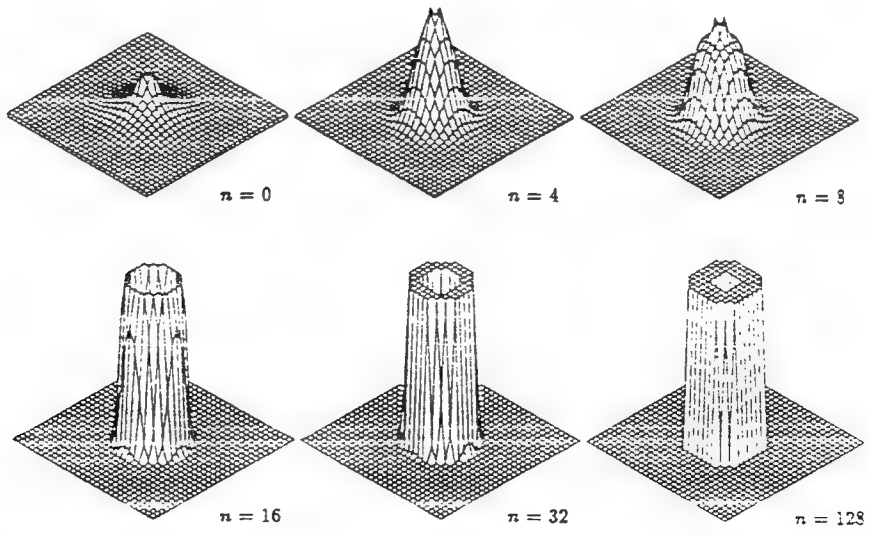



**Figure 9.** The reconstructed imaginary values of the contrast for example 1. The maximum reconstructed contrast is constrained to 5.97.

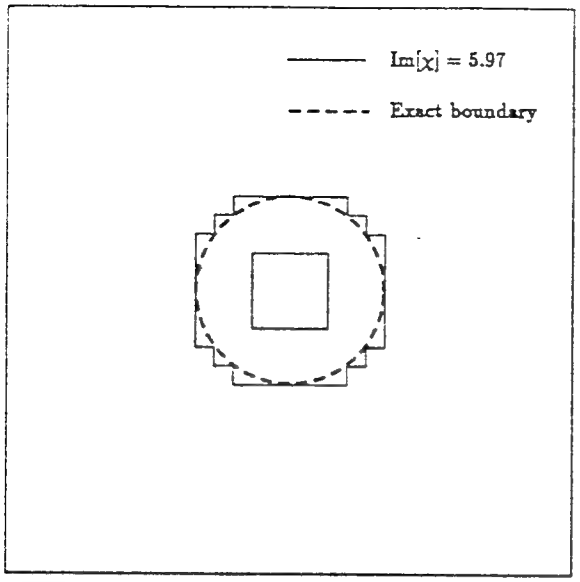



**Figure 10.** Comparison between the reconstructed boundary of example 1 ($n = 128$) and the exact one.

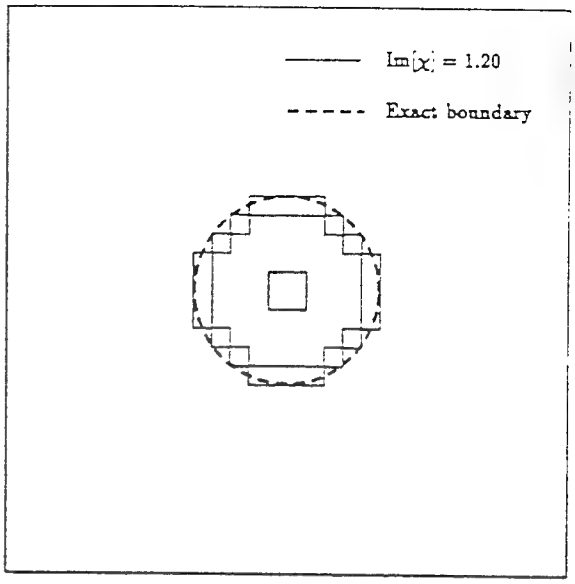



**Figure 12.** Comparison between the reconstructed boundary of example 2 ($n = 64$) and the exact one.

by the dashed circle. The outer contour line approximates the boundary very well.

### Example 3

The reconstruction of the shifted cylinder is shown in Figure 13. It again shows that our scheme not only approximates the boundary of the object very well, but also the location is determined precisely. This is stressed in Figure 14, where after 64 iterations the contour lines $\chi = \chi_{max}$ have been plotted. Again, the exact boundary of the object is indicated by the dashed circle.

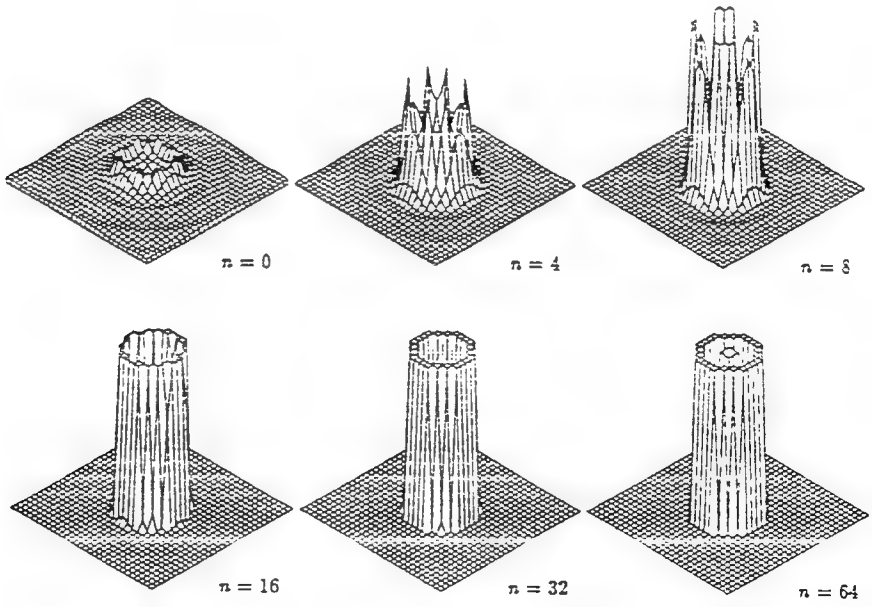



**Figure 11.** The reconstructed imaginary values of the contrast for example 2. The maximum reconstructed contrast is constrained to 1.20.

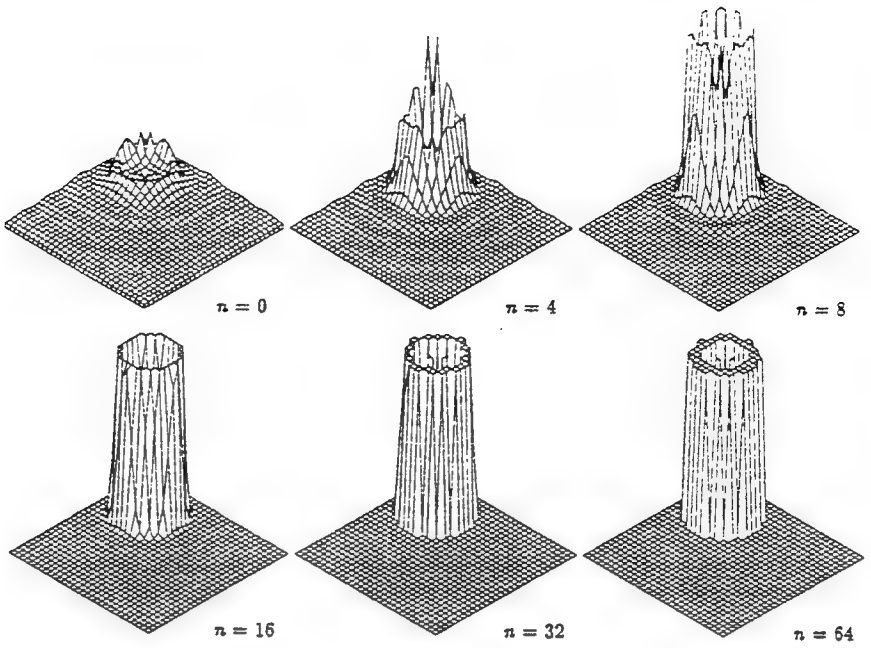



**Figure 13.** The reconstructed imaginary values of the contrast for example 3. The maximum reconstructed contrast is constrained to 1.20.

For this example we also investigate the influence of noisy data. We have added to the data a random noise signal with maximum amplitude of 50% of the maximum amplitude of the data. The reconstruction process is shown in Figure 15. It is observed that this extremely high noise level yields some local anomalies, but the location and shape of the cylinder is still clearly visible in the reconstructed contrast. This example indicates the robustness of our reconstruction scheme.

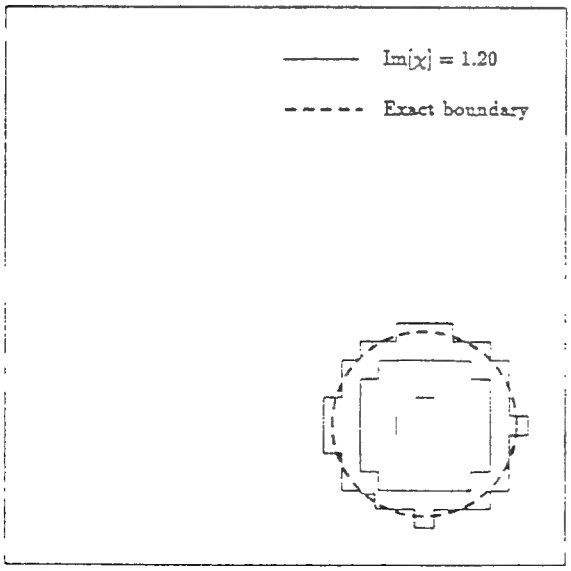



**Figure 14.** Comparison between the reconstructed boundary of example 1 ($n = 64$) and the exact one.

The computer code was run on a VAX-4000 workstation. The last example requires about 8 Mbyte memory, while one iteration takes one minute CPU time.

### Finer Mesh

Finally, we present the reconstruction of the third example when the test domain is subdivided into a finer mesh. Now the test square is subdivided into $61 \times 61$ subsquares of $0.0015 \times 0.0015$ m$^2$. The reconstruction is shown in Figures 16 and 17.

## Conclusions

An iterative method for reconstructing complex constitutive parameters has been modified to reconstruct the location and shape of impenetrable objects by exploiting the fact that, electromagnetically impenetrable objects are really lossy dielectrics with very high conductivity so that the skin depth is very small, hence the data from impenetrable scatterers is consistent with the re-

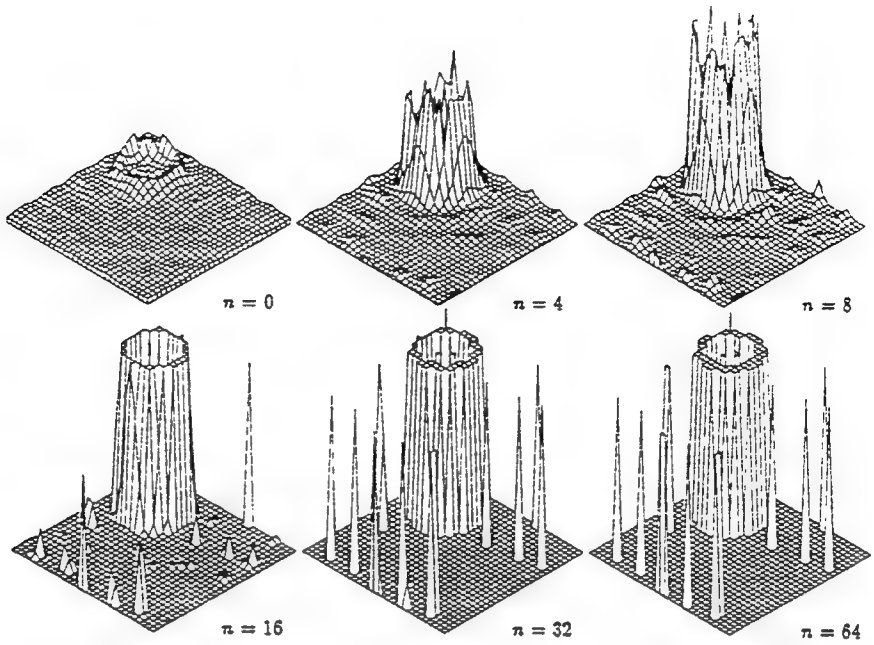



**Figure 15.** The reconstructed imaginary values of the contrast for example 3 from data with 50% noise. The maximum reconstructed contrast is constrained to 1.20.

construction algorithm. Since the incident field decays drastically as it penetrates the body, the only reliable information about the body that can be inferred from scattered field data comes from a neighborhood of the surface. Using this fact, we employ an algorithm designed to reconstruct the conductivity (and permittivity) throughout the body but give credence only to the boundary of the support of the reconstructed conductivity when

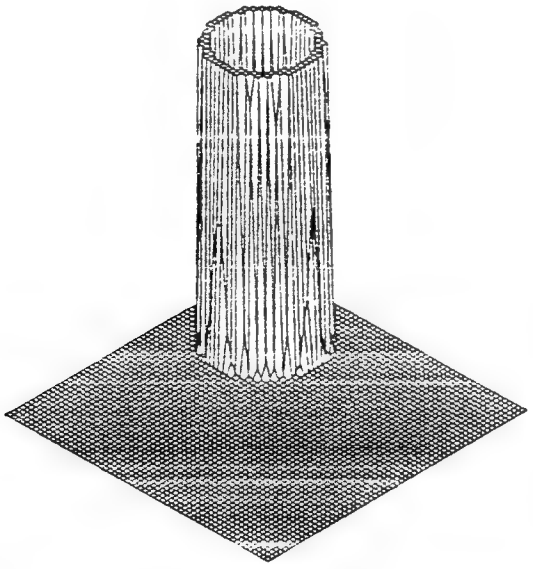


**Figure 16.** The reconstructed imaginary values of the contrast for example 3 and a refined mesh. The maximum reconstructed contrast is constrained to 3.09.

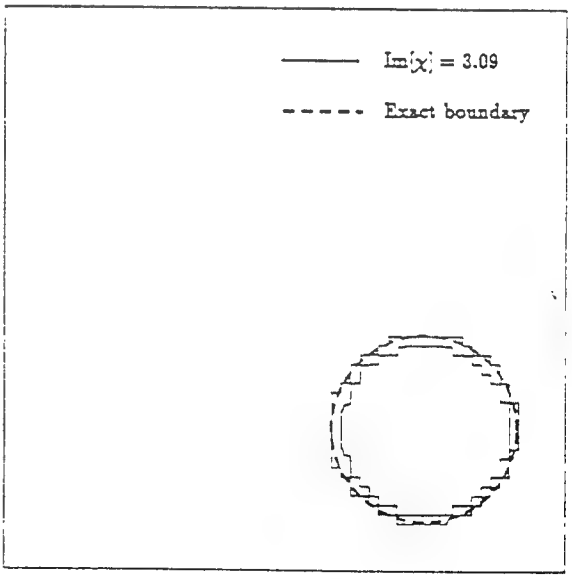



**Figure 17.** Comparison between the reconstructed boundary of example 3 ($n = 64$) and the exact one.

these values are large. Numerical evidence is presented which shows the utility of this approach. A perfectly conducting circular cylinder was taken as a target for $ka = \pi/3$ and $ka = \pi$. The synthetic data were obtained from the exact solution available through the use of separation of variables, and examples were treated for the cylinder both centered and off centered in the test domain. The effect of noise was examined as was the result of refining the discretization of the test domain. The numerical examples show that the algorithm is effective in confining the boundary to an annular domain, and the resolution increases as the wavelength decreases and also as the discretization of the test domain becomes finer. Moreover, the results appear to be remarkably stable with respect to noise.

The numerical results reported here concern only circular cylindrical scatterers, although the origin was shifted so as to remove some of the effects of symmetry. However, the method applies equally to noncircular objects and has been successful in reconstructing a noncircular scatterer from experimental data. These results will be reported elsewhere. It should also be noted that although the results reported here involved far-field data, additional experiments using near-field data yielded comparable results. That is, the reconstructions were of the same quality with the same number of iterations.

Ongoing work is concerned with further refinements to make the algorithm more efficient by successively reducing the number of points or subsquares at which the contrast is updated once very large or very small values are attained. Also under investigation are the improvements resulting when multifrequency data are used and the performance of the algorithm when real rather than synthetic data are employed. Results of these studies will be reported in the future.

Acknowledgments. This work was supported under AFOSR grant 91-0277 and NATO grant 0230/88.

## References

Angell, T. S., R. E. Kleinman, and G. F. Roach, An inverse transmission problem for the Helmholtz equation, *Inverse Probl.*, *3*, 149–180, 1986.

Angell, T. S., R. E. Kleinman, B. Kok, and G. F. Roach, A constructive method for identification of an impenetrable scatterer, *Wave Motion*, *11*, 185–200, 1989.

Brodlie, K. W., Unconstrained minimization, in *The State of the Art in Numerical Analysis*, edited by D. A. H. Jacobs, pp. 255–257, Academic, San Diego, Calif., 1977.

Colton, D., and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer-Verlag, New York, 1992.

Colton, D., and P. Monk, The numerical solution of the three dimensional inverse scattering problem for time harmonic acoustic waves, *SIAM J. Sci. Stat. Comput.*, *8*, 278–291, 1987.

Habashy, T. M., M. L. Oristaglio, and A. T. de Hoop, Simultaneous inversion of permittivity and conductivity employing a nonpertubative approach, *SPIE 1767*, *Inverse Problems in Scattering and Imaging*, 193–205, 1992.

Habashy, T. M., M. L. Oristaglio, and A. T. de Hoop, Simultaneous nonlinear reconstruction of two-dimensional permittivity and conductivity, *Radio Sci.*, this issue, 1994.

Kleinman, R. E., and P. M. van den Berg, A modified gradient method for two-dimensional problems in tomography, *J. Comput. Appl. Math.*, *42*, 17–35, 1992.

Kleinman, R. E., and P. M. van den Berg, An extended range modified gradient technique for profile inversion, *Radio Sci.*, *28*, 877–884, 1993.

Kohn, R. V., and A. McKenney, Numerical implementation of a variational method for electrical impedance tomography, *Inverse Probl.*, *6*, 389–414, 1990.

Press, W. H., B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes. The Art of Scientific Computing*, Cambridge University Press, New York, 1986.

Richmond, J. H., Scattering by a dielectric cylinder of arbitrary cross section shape, *IEEE Trans. Antennas Propag.*, *AP-13*, 334–341, 1965.

Sabbagh, H. A., and R. G. Lautzenheiser, Inverse problems in electromagnetic nondestructive evaluation, *Int. J. Appl. Electromagn. Mater.*, *3*, 253–261, 1993.

Van den Berg, P. M., Iterative computational techniques in scattering based upon the integrated square error criterion, *IEEE Trans. Antennas Propag.*, *AP-32*, 1063–1071, 1984.

Wexler, A., B. Fry, and M. R. Neuman, Impedance-computed tomography algorithm and system, *Appl. Opt.*, *24*, 3985–3992, 1985.

R. E. Kleinman, Center for the Mathematics of Waves, Department of Mathematical Sciences, University of Delaware, Newark, DE 19716.

P. M. van den Berg, Laboratory of Electromagnetic Research, Faculty of Electrical Engineering, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, Netherlands.

(Received July 23, 1993; revised November 8, 1993; accepted November 8, 1993.)

SECOND INTERNATIONAL CONFERENCE ON

# MATHEMATICAL AND NUMERICAL ASPECTS OF WAVE PROPAGATION

Edited by Ralph Kleinman
Thomas Angell
David Colton
Fadil Santosa
Ivar Stakgold
University of Delaware

siam. Philadelphia
Society for Industrial and Applied Mathematics

# Chapter 31
# Full Low-Frequency Asymptotic Expansion for Elliptic Equations of Second Order

R. Kleinman† B. Vainberg‡

**Abstract**

The present paper shows how to obtain the low frequency expansions of solutions of a large class of exterior boundary value problems involving second order elliptic equations in two dimensions. The differential equations must coincide with the Helmholtz equation in a neighborhood of infinity, however they may depart radically from the Helmholtz equation in any bounded region provided they retain ellipticity. The procedure for determining the full low frequency expansion of solutions of the exterior Dirichlet and Neumann problems for the Helmholtz equation is included as a special case of the results presented here.

## 1 Introduction and Formulation of the Main Results

Let $\Omega$ be an unbounded domain in $R^2$ with compact infinitely smooth boundary $\Gamma$, let

$$A = \sum_{i,j=1}^{2} \frac{\partial}{\partial x_i} a_{ij}(x) \frac{\partial}{\partial x_j} + \sum_{i=1}^{2} b_i(x) \frac{\partial}{\partial x_i} + c(x)$$

be an elliptic operator of the second order (that is the matrix $(a_{ij}(x))$ is non-singular) with infinitely smooth coefficients in $\bar{\Omega}$ and $a_{ij}(x)$ real valued and let $A$ coincide with the Laplace operator $\Delta$ in some neighborhood of infinity. Denote by $u$, a solution of the problem

$$\begin{cases} Au + k^2 u = f, & x \in \Omega \\ Bu = 0, & x \in \Gamma \end{cases} \tag{1}$$

where $B$ is either the identity (Dirichlet boundary condition) or the following operator

$$Bu = \frac{\partial u}{\partial \nu} + p(x) \frac{\partial u}{\partial l} + q(x)u. \tag{2}$$

Here $\frac{\partial}{\partial \nu} = \sum_{i,j=1}^{2} a_{ij}(x) n_i \frac{\partial}{\partial x_j}$ is the derivative along the conormal vector ($n = (n_1, n_2)$ is the unit vector which is normal to $\Gamma$ and directed into $\Omega$), $\frac{\partial}{\partial l}$ is the derivative along $\Gamma$, $p, q \in C^\infty$. Finally denote by

$$R_k u : L_2(\Omega) \to H^2(\Omega), \quad \operatorname{Im} k > 0$$

† University of Delaware, Newark, Delaware 19716

‡ University of North Carolina, Charlotte, North Carolina 28223

the operator which takes functions $f \in L_2(\Omega)$ into solutions of problem (1) belonging to the Sobolev space $H^2(\Omega)$.

With $a$ an arbitrary constant we define a cutoff function $\chi = \chi(x) \in C^\infty(\bar{\Omega})$, such that $\chi = 1$ when $|x| < a$ and $\chi = 0$ when $|x| > a + 1$. Then a restricted resolvent is defined as

$$\hat{R}_k := \chi R_k \chi : L_2(\Omega) \to H^2(\Omega).$$

The operators $\hat{R}_k, R_k$ are defined and are meromorphic functions of $k$ when $\operatorname{Im} k > 0$. Moreover the operator $\hat{R}_k, \operatorname{Im} k > 0$, has a meromorphic continuation on the Riemann surface of the function $\ln k$(see [1]). Let us stress that $\hat{R}_k f = \chi R_k f$ if $f = 0$ for $|x| > a$. In this case the function $u = \hat{R}_k f$ is a solution of (1) for $|x| < a$.

The present work is devoted to the study of the asymptotic behavior of the operator $\hat{R}_k$ (that is of the solution $u = \hat{R}_k f$, $|x| < a$ of the problem (1) with $f = 0$ for $|x| > a$) as $k \to 0$. We consider only the two-dimensional case, since in other dimensions the asymptotic behavior of the solution of this problem is much simpler. The most recent results for the two-dimensional case as well as extensive references to earlier work are found in [2]. There the problem was supposed to be formally self-adjoint (i.e. $(Au, v) = (u, Av)$ for functions $u, v \in C_0^\infty(\bar{\Omega})$ satisfying specified boundary conditions) and nonpositive, or to be more exact it was supposed that

$$(Au, u) \le -\alpha \int_\Omega |\nabla u|^2 dx, \ \alpha > 0 \quad \text{for all } u \in C_0^\infty(\bar{\Omega}) \text{ with } Bu|_\Gamma = 0. \tag{3}$$

Here $(\cdot, \cdot)$ denotes the inner product in $L_2(\Omega)$. For this case [2] gives the asymptotic behavior of the solution $u = u_k$ of the problem (1) with accuracy $O(k^2)$.

However more than ten years ago, in [1], there appeared results of one of the present authors concerning the low-frequency asymptotic behavior of solutions of general elliptic problems of any order polynomially depending on the spectral parameter. Those results apply to problem (1) and allow one to obtain the full asymptotic expansion of the operator $\hat{R}_k$ as $|k| \to 0$. This expansion has the form

$$\hat{R}_k = k^{-\alpha} \sum_{m=0}^{\infty} \sum_{n=0}^{m\ell} \frac{k^m}{P(\ln k)^m} \ln^n k \, P_{m,n}, \quad |k| << 1, \tag{4}$$

where $\alpha$ is an integer, $\ell$ is a non-negative integer, $P$ is a polynomial with constant coefficients, and $P_{m,n} : L_2(\Omega) \to H^2(\Omega)$ are bounded operators independent of $k$.

The integers $\alpha$ and $\ell$ and the polynomial $P$ are not known in general, even for equations of 2nd order. It is the purpose of the present work to specify the precise form of expansion (4) for solutions of two restricted cases of problem (1).

Case I. The space of bounded solutions of the homogeneous problem

$$Au = 0, \ x \in \Omega; \ Bu = 0, \ x \in \Gamma \tag{5}$$

consists of only the trivial solution.

Case II. The space of bounded solutions of (5) is one dimensional and if $u$ is a nontrivial solution then

$$\lim_{r \to \infty} u(x) \neq 0. \tag{6}$$

and the formal adjoint to problem (5) (see (13) below) also has a bounded solution with property (6).

The exterior Dirichlet problem for the Helmholtz equation is an example of Case I while the exterior Neumann problem for the Helmholtz equation is an example of Case II.

It is well known that if $u$ is a bounded solution of Laplace's equation in a neighborhood of $\infty$ then for $r$ sufficiently large, $u$ has the form

$$u(x) = C_0 + \sum_{n=0}^{\infty} (a_n \cos n\varphi + b_n \sin n\varphi) r^{-n} \tag{7}$$

In particular this representation is valid for bounded solutions of problem (5) and of problem (1) with $k = 0$ if $f = 0$ in a neighborhood of infinity. Therefore, condition(6) is equivalent to the requirement that $C_0 \neq 0$ for such solutions.

The conditions embodied in Cases I and II are less restrictive than those used in [2]. If condition (3) required in [2] is fulfilled then together with (7) it follows that there are only constant solutions of problem (5). This means that either Case I or Case II apply. In our work we do not require nonpositivity, condition (3), nor do we require that the problem be self adjoint. Moreover we obtain not only the first few terms as in [2], but the complete asymptotic expansion of the solutions.

Unlike [2] we consider (for simplicity) only the problems in which the boundary and coefficients of the equation are infinitely smooth.

The main results are contained in two theorems which are presented in this paper. Let $a$ be an arbitrary fixed constant such that $\Gamma$ is contained in the circle $|x| < a - 1$ and $f = 0$ when $|x| > a$. Let $\Omega_a = \Omega \cap \{x : |x| < a\}$ and $A = \Delta + k^2$ when $|x| > a - 1$. Let $L_{2,a}$ be the space of functions which belong to $L_2(\Omega)$ and are equal to zero when $|x| > a$. In particular, $f \in L_{2,a}$.

In Case I we denote by $u_0, u_1, u_2$ the solutions of the problems

$$\begin{cases} Au_0 = f, & x \in \Omega \\ Bu_0 = 0, \ x \in \Gamma; & |u_0| < \infty \text{ as } r \to \infty \end{cases} \tag{8}$$

$$\begin{cases} Au_1 = 0, & x \in \Omega; \\ Bu_1 = 0, \ x \in \Gamma; & |u_1 - \ln r| < \infty \text{ as } r \to \infty \end{cases} \tag{9}$$

$$\begin{cases} Au_2 = 0, & x \in \Omega \\ Bu_2 = 0, \ x \in \Gamma; & |u_2 + \frac{1}{2}a_1 x_1 + \frac{1}{2}b_1 x_2| < \infty \text{ as } r \to \infty \end{cases} \tag{10}$$

where $a_1, b_1$ are the coefficients in the expansion (7) for $u_0$. We show that the uniqueness of the solution of problem (8) leads to the solvability of this problem. After we have established this, we can easily infer the unique solvability of problems (9) and (10) by reducing them to problems of the form (8). This is accomplished by writing the solutions $u_1$ and $u_2$ of problems (9) and (10) in the form $u_1 = v \ln r + w_1$ and $u_2 = -\frac{1}{2}(a_1 x_1 + b_1 x_2)v + w_2$, where $v = v(x)$ is a cutoff function which is equal to one for $|x| > c$ and equal to zero in some neighborhood of $\Gamma$. Then the problem of finding $w_i$ is of the form (8). Moreover since the expansion (7) is valid for $w_i$, in particular the constant

$$\lambda_0 = \lim_{r \to \infty} (u_1 - \ln r) \tag{11}$$

is defined. Finally, we denote by $\beta$ the constant which occurs in the asymptotic expansion of $H_0^{(1)}(z)$, the Hankel function of the first kind and order zero:

$$H_0^{(1)}(z) = \frac{2i}{\pi}(\ln z - \beta) + O(z^2 \ln z), \ z \to 0.$$

With this notation established we now state the main results.

THEOREM 1. *In Case I for the solution $u = \hat{R}_k f$ of problem (1) with $f \in L_{2,a}$, the following asymptotic expansion is valid when $-\frac{\pi}{2} \le arg k \le \frac{3\pi}{2}$, $|k| \to 0$:*

$$u = \sum_{m=0}^{N} \sum_{n=0}^{m} \sum_{p=0}^{m-n+1} k^{2m} ln^n k (lnk - \lambda_0 - \beta)^{-p} u_{m,n,p}(x) + \tilde{u}_N \tag{12}$$

*where $u_{m,n,p}(x)$ are independent of $k$ and*

$$\|\tilde{u}_N\|_{H^2(\Omega_a)} \le C(a)|k^2 lnk|^{N+1} \|f\|_{L_{2,a}}.$$

*The leading terms of the asymptotic expansion have the form*

$$u = u_0(x) + \frac{C_0}{lnk - \lambda_0 - \beta} u_1(x) + k^2 lnk u_2(x) + O(k^2)$$

*where $u_0, u_1, u_2$ are the solutions of problems (8) - (10) and $C_0 = \lim_{r \to \infty} u_0(x)$.*

**Remark:** In fact the corresponding expansion for the operator $\hat{R}_k$ converges in the operator norm for $0 < |k| < |k_0|$ for some $|k_0| > 0$ and therefore the infinite series for $u (N = \infty)$ converges in $H^2(\Omega_a)$.

In Case II let us denote by $v_0$ the solution of problem (5) such that

$$\lim_{r \to \infty} v_0(x) = 1.$$

Let the problem

$$A^* u = 0, \; x \in \Omega; \quad B^* u = 0, \quad x \in \Gamma \tag{13}$$

be formally adjoint to (5), that is, the operator $A^*$ can be obtained from $A$ by substituting $\bar{b}_i$ for $b_i$ and $\bar{c} - \sum \frac{\partial \bar{b}_i}{\partial x_i}$ for $c$. If $B = I$ then $B^* = I$, if $B$ has the form (2) then $B^*$ has the same form with $\bar{p}$ instead of $p$ and $-\frac{\partial \bar{p}}{\partial t} - \bar{q} - \sum \bar{b}_i(x) n_i$ instead of $q$. If $u, v \in C^\infty(\bar{\Omega})$, and $Bu = B^* v = 0$ on $\Gamma$ then

$$\int_{\Omega_R} Au\bar{v}dx = \int_{\Omega_R} u\bar{A}^* v dx + \int_{r=R} (\frac{\partial u}{\partial r}\bar{v} - u\frac{\partial \bar{v}}{\partial r}) dS, \; R > a.$$

We show that the space of bounded solutions of problem (13) in Case II is also one-dimensional and there exists a unique solution $v_*$ of problem (13) such that

$$\lim_{r \to \infty} v_*(x) = 1.$$

Let us denote by $v_1$ the solution of the inhomogeneous problem (5) (that is the solution of problem (1) with $k = 0$) such that

$$v_1 - \alpha(\ln r - \beta) \to 0 \text{ as } r \to \infty$$

where $\alpha = \alpha(f)$ is constant. We show that in Case II such a solution exists, is unique and

$$\alpha = \frac{1}{2\pi} \int_{\Omega} f\bar{v}_* dx \tag{14}$$

THEOREM 2. *In Case II for the solution $u = \dot{R}_k f$ of problem (1) with $f \in L_{2,a}$, the following asymptotic expansion is valid when $-\frac{\pi}{2} \le \arg k \le \frac{3\pi}{2}$, $|k| \to 0$:*

$$u = \sum_{m=0}^{N} \sum_{n=0}^{2m+1} k^{2m} \ln^n k \; u_{m,n}(x) + \tilde{u}_N \tag{15}$$

*where $u_{m,n}$ are independent of $k$ and*

$$\|\tilde{u}_N\|_{H^2(\Omega_a)} \le C(a)|k^{2N+2} \ln^{2N+3} k| \|f\|_{L_{2,a}}$$

*The leading terms of the asymptotic expansion have the form*

$$u = \alpha \ln k v_0(x) + v_1(x) + O(k^2 \ln^3 k), \quad |k| \to 0$$

*where $\alpha$ is defined in (14).*

*If $v_0 \equiv 1$ then $u_{mn} \equiv 0$ for $n > m + 1$.*

**Remarks.**

(1) The remark following Theorem I also applies here.

(2) It is not difficult to write out the sequence of problems similar to (8) - (10), from which we can find all the coefficients in the expansions (12) and (15).

(3) In the present work we have assumed that the data $f$ is independent of $k$. In many applications, of course, $f$ will be a known function of $k$. In such cases when $f$ may developed in a series in $k$ the present analysis will still apply. The result will be the product of the expansion of the inverse operator, $\dot{R}_k$, with the expansion of $f$.

Theorems 1 and 2 are proved in a similar fashion and the details will be presented elsewhere. Here we will provide an outline of the main ideas in the proof of Theorem 1.

It consists first of establishing the expansion

$$u = u_0 + \frac{C_0}{\ln k - \lambda_0 - \beta} v_1(x) + O(k^2 \ln^2 k), \; x \in \Omega_a, \; k \to 0 \tag{16}$$

for the solution $u = \dot{R}_k f$ of problem (1) with $f \in L_{2,a}$ and $u_0, u_1, C_0, \lambda_0$ and $\beta$ as in Theorem 1. This expansion is obtained from the asymptotic expansion of the resolvent in (4), the differential equation (1) and an integral representation of the solution (actually an integral equation) based on Green's theorem, namely

$$\eta(x)u(x) = \frac{1}{2\pi} \int_\Omega (\ln k + \ln|x-y| - \beta + O(k^2 \ln k))(u(y)\Delta_y \eta + 2\nabla_y u \cdot \nabla_y \eta) dy, \; x \in \mathbb{R}^2, \; k \to 0 \tag{17}$$

where $\eta \in C^\infty(\mathbb{R}^2)$ and $\eta = 0$, $|x| < a - 1$, $\eta = 1$, $|x| > a - \frac{1}{2}$.

The last step of the proof is based on a special parametrix of problem (1). This parametrix involves the operater $U$ which is defined as that operator which maps any function $f \in L_{2,a}$ into the sum of the first two terms on the right hand side of equation (16). Then choosing $\zeta \in C^\infty(\mathbb{R}^2)$ such that $\zeta = 1$ for $|x| > a - 1$ and $\zeta = 0$ in a neighborhood of $\Gamma$ we define the operator $\Phi_k$ as

$$\Phi_k h = (1 - \eta)Uh - \frac{i\zeta}{4} \int_\Omega H_0^{(1)}(k|x-y|)\Delta_y(\eta U h) dy. \tag{18}$$

Now we assume a solution of problem (1) in the form $u = \Phi_k h$ with unknown $h \in L_{2,a}$. This leads to the equation $(I + T_k)h = f$ for $h$, where the norm of $T_k$ is small by virtue of

equation (16). This allows us to obtain an asymptotic expansion for $(I+T_k)^{-1}$ and then for $\hat{R}_k = \Phi_k(I+T_k)^{-1}$.

Acknowledgment: This work was partially supported under AFOSR Grant 91-0277.

## References

[1] B. R. Vainberg, *Asymptotic Methods in Equations of Mathematical Physics*, Moscow State Univ. Publishers, 1982, (in Russian); English Translation (1989), Gordon and Breach Publishers.

[2] H. Weck and K. J. Witsch, *Exact Low Frequency Analysis for a Class of Exterior Boundary Value Problems for the Reduced Wave Equations in Two Dimensions*, SonderForschungsbereich 256, Report 117, Bonn University, 1990.

# ASYMPTOTIC APPROXIMATION OF OPTIMAL SOLUTIONS OF AN ACOUSTIC RADIATION PROBLEM

**T. S. Angell and R. E. Kleinman**
Center for the Mathematics of Waves
Department of Mathematical Sciences
University of Delaware
Newark, DE 19716

and

**B. Vainberg**
Department of Mathematical Sciences
University of North Carolina
Charlotte, NC 28223 .

**Technical Report No. 93–10**

To appear in: Inverse Scattering and Potential Problems in Mathematical Physics, R. E. Kleinman, R. Kress, and E. Martensen, eds., Peter Lang, Frankfurt.

# Asymptotic Approximation of Optimal Solutions of an Acoustic Radiation Problem[1]

**T. S. Angell and R. E. Kleinman**
Center for the Mathematics of Waves
Department of Mathematical Sciences
University of Delaware
Newark, DE 19716
and
**B. Vainberg**
Department of Mathematical Sciences
University of North Carolina
Charlotte, NC 28223

**Abstract** We have presented, elsewhere, the problem of choosing Neumann data for the exterior Helmholtz equation in order to optimize a functional of the radiated far field. In this paper we use asymptotic methods to determine an approximate optimal solution whose support is in a prescribed region of the boundary.

## I. Introduction

Let $\Omega$ be exterior of a strongly convex bounded obstacle $B \subset I\!R^n$ with infinitely smooth boundary $\Gamma$. Let $u$ be a solution of the problem

$$(1) \qquad \begin{cases} (\Delta + k^2)u = 0, & x \in \Omega \\ \frac{\partial u}{\partial n} = h, & x \in \Gamma \\ \frac{\partial u}{\partial r} - iku = o(r^{\frac{1-n}{2}}), & r \to \infty \end{cases}$$

where $h$ is an infinitely smooth function and $\partial/\partial n$ is the derivative in the direction of the (exterior) normal which is directed into the unbounded region.

[1] AFOSR # 86–0269

It is well known that problem (1) is uniquely solvable. Let us denote by $T$ the Neumann-to-Dirichlet operator, which transforms $h$ into $u|_\Gamma$, where $u$ is the solution of problem (1). This operator, initially defined on smooth functions, can be extended as a bounded operator on the whole space $L^2(\Gamma)$ since $T$ is a pseudodifferential operator of order $-1$ [9, 10]. Let us introduce the following norm in the space $L^2(\Gamma)$:

$$|||h||| = [k^{-2}\|h\|^2_{L^2(\Gamma)} + \|Th\|^2_{L^2(\Gamma)}]^{1/2}$$

It is also well known that the solution $u$ of problem (1) has the asymptotic behavior in the far-field:

$$u = f\left(\frac{x}{|x|}, k\right) \frac{e^{ikr}}{r^{\frac{n-1}{2}}}(1 + 0(r^{-1})),\ r \to \infty$$

where $f(\theta, k),\ \theta \in S^{n-1}$, is a smooth function which has the form

$$f(\theta, k) = \beta_n \int_\Gamma [h + ik < \theta, \hat{\mathrm{n}} > Th] e^{-ik<\theta, y>} dS_y. \tag{2}$$

Here

$$\beta_n = \beta_n(k) = -\frac{1}{4\pi}\left(\frac{k}{2\pi i}\right)^{(n-3)/2}. \tag{3}$$

Let $a = a(\theta)$ be a piecewise continuous non-negative function on the unit sphere. Let the functional $F$ be defined by

$$F(h) = \int_{S^{n-1}} |f(\theta)|^2 a(\theta) dS,\ \ h \in L^2(\Gamma),$$

where $S$ is the element of surface area.

We are interested in the maximum value of the functional $F$ on the set $U$ of functions $h$ in $L^2(\Gamma)$ with $|||h||| = 1$. In addition we are interested in characterizing

the functions $h$, $|||h||| = 1$, where $F$ attains its maximum. The existence of such functions $h$ follows from the results of [1], [2].

In order to formulate the main result we need to introduce some notation. Let the mapping $P : \Gamma \to S^{n-1}$ transfer each point $x \in \Gamma$ into the point $\theta \in S^{n-1}$ for which $\hat{\mathbf{n}} = \theta$, where $\hat{\mathbf{n}}$ is the unit vector of exterior normal to $\Gamma$ at the point $x$. Note that we will use the symbol $\theta$ to denote both a point on the unit sphere and the position vector of that point. For an arbitrary $\epsilon > 0$ we construct a function $g_\epsilon = g_\epsilon(\theta)$ such that

$$\int_{S^{n-1}} |g_\epsilon(\theta)|^2 dS = 1, \quad \int_{S^{n-1}} |g_\epsilon(\theta)|^2 a(\theta) dS \geq \sup a(\theta) - \epsilon. \tag{4}$$

It is obvious that we can take

$$g_\epsilon = \varphi / \Big( \int_{S^{n-1}} |\varphi|^2 dS \Big)^{1/2}$$

where $\varphi$ is any function on $S^{n-1}$ with support in a region where $a(\theta) > \sup a(\theta) - \epsilon$. Let $\kappa(x)$ be the total curvature (product of the principle curvatures) of $\Gamma$ at the point $x \in \Gamma$.

The main result of the paper is contained in the following theorem.

**Main Theorem.** 1. If $|||h||| = 1$ then

$$0 \leq F(h) \leq \frac{1}{2} \sup a(\theta). \tag{5}$$

2. Let $\epsilon > 0$ be an arbitrary positive number, $g_\epsilon$ be a fixed function (independent of $k$) which satisfies the relations (4) and

$$h_\epsilon = h_\epsilon(x) = \frac{\sqrt{2}k}{2} g_\epsilon(Px) \sqrt{\kappa(x)}. \tag{6}$$

Then

$$|||h_\epsilon||| = 1 + 0(k^{-1}),\ n \to \infty \tag{7}$$

and there exist $k_0 = k_0(\epsilon)$ such that

$$F(h_\epsilon) > \frac{1}{2}\max a(\theta) - 2\epsilon \tag{8}$$

if $k \geq k_0$.

From this theorem it follows that if

$$\tilde{h}_\epsilon = h_\epsilon / |||h_\epsilon|||$$

then $F(\tilde{h}_\epsilon)$ differs from its maximum value on the set $U = \{h \in L^2(\Gamma) : |||h||| = 1\}$ by not more than $3\epsilon$ if $k$ is sufficiently large.

## II. Asymptotic Behavior of the Solutions of the Problem (1)

**Theorem 1.** If $h$ is independent of $k$ then there exist infinitely smooth functions $a_j(x)$ such that the solution of the problem (1) has the following asymptotic expansion:

$$u = e^{ikS(x)}[\sum_{j=0}^{N} a_j(x)(ik)^{-j-1} + u_N] \tag{9}$$

where $S(x)$ is the distance between a point $x$ and $\Gamma$, $a_j \in C^\infty(\bar{\Omega})$, $a_0(x) = h(x)$ on $\Gamma$ and

$$|\partial_x^\alpha u_N| \leq C k^{-N-2},\ |x| \leq a,\ k \geq 1$$

for any $a < \infty$, $\alpha = (\alpha_1, \ldots, \alpha_n)$ and some constant $C = C(a, \alpha, h)$.

**Corollary**: If $x \in \Gamma$ then

$$u = (ik)^{-1} h(x) + k^{-2} u_1$$

where the function $u_1 = u_1(x,k)$ and any of its derivatives along $\Gamma$ are bounded when $k \geq 1$.

This result is obvious from the point of view of physics. But the strict mathematical proof is not very simple, as the problem under consideration involves two large parameters: as $|x| \to \infty$, the unique solution is singled out by (radiation) conditions, and, as $k \to \infty$, we are interested in the asymptotic behavior of the solution. High frequency asymptotic results have been obtained previously for the problem of scattering of plane waves by an inhomogeneous medium [3],[4] and by obstacles [5],[6]. However our results are much simpler due to the fact that in the present cases $h$ is independent of $k$ (or has a "simple" $k$ dependence) so that no caustics occur. In order to prove Theorem 1, we will follow the same technique introduced in [4] and employ the nonstationary problem, corresponding to Problem (1):

$$\begin{cases} v_{tt} - \Delta v = 0, & x \in \Omega,\ t \in I\!R; \\ \frac{\partial v}{\partial n}\big|_\Gamma = h\beta(t)e^{-ikt},\ t \in I\!R; & v \equiv 0,\ t < 0, \end{cases} \tag{10}$$

where $\beta \in C^\infty$, $\beta(t) = 0$ when $t < 1/2$, $\beta(t) = 1$ when $t > 1$.

The connection between problems (1) and (10) is established by Theorem 3, below, the *principle of limiting amplitude*. However, we will need more accurate estimates of the remainder than those which are usually used in that principle. For this we need the following uniform estimate of the solutions of the initial-boundary value problem with homogeneous boundary data

$$\begin{cases} w_{tt} - \Delta w = 0, & x \in \Omega,\ t > 0; \\ \frac{\partial w}{\partial n}\big|_\Gamma = 0; & w\big|_{t=0} = 0,\ w'_t\big|_{t=0} = f(x) \end{cases} \tag{11}$$

where $f \in L^2(\Omega)$, $f(x) = 0$ when $|x| > a$.

**Theorem 2** For any $a$ there exists $T_0 = T_0(a)$ such that the following estimates are satisfied

$$|\partial_t^j \partial_x^\alpha w| \le C|\partial_t^j \gamma(t)|\|f\|_{L^2},\ t > T_0,\ |x| < a \tag{12}$$

where $j$ and $\alpha = (\alpha, \ldots, \alpha_n)$ are arbitrary (non-negative integers), $C$ depends on $a$, $j$ and $\alpha$ but not on $f$ and

$$\gamma(t) = \begin{cases} e^{-\epsilon t},\ \epsilon > 0, & \text{if } n \text{ is odd} \\ t^{2-n} \ln t & \text{if } n \text{ is even} \end{cases}$$

This large time behavior of $w$ may be obtained from Theorems 4 and 6 of Chapter 10 in [4].

**Remark:** This theorem shows that the solution and all derivatives decay as $t \to \infty$ uniformly with respect to the initial data if $n > 2$. If $n = 2$ it is possible to show that $(w - \tilde{c}_1 \ln t - \tilde{c}_2)$ will decay for some constants $\tilde{c}_1, \tilde{c}_2$.

Now we establish the *limiting amplitude principle* in the form which we need for the proof of Theorem 1. Let $\varphi \in C^\infty(\mathbb{R})$, $\varphi(t) = 0$ when $t < 0$ or $t > 1$, $\int_0^1 \varphi(t)dt = 1$.

**Theorem 3.** The solutions of problem (10) can be represented in the form of

$$v = u(x)e^{-ikt} + v_1 \tag{13}$$

where $u$ is the solution of problem (1) and for any $a < \infty$, $T > T_0(a) + 1$, any $N$ and $\alpha = (\alpha_1, \ldots, \alpha_n)$, the following estimates are valid for $v_1$

$$\|\partial_x^\alpha \int_T^{T+1} v_1 e^{ikt} \varphi(t - T)dt\| \le Ck^{-N},\ |x| \le a,\ k \ge 1 \tag{14}$$

with constant $C = C(a, T, N, \alpha, h)$ which does not depend on $k$.

**Proof.** With no loss of generality we can assume that $\partial\Omega \subset \{x : |x| < a\}$. Let $\psi$ be a function such that $\psi \in C^\infty(\bar{\Omega})$, $\psi = 0$ when $|x| > a$, $\frac{\partial\psi}{\partial n} = h$ on $\Gamma$. It is obvious that the function

$$v = \psi(x)\beta(t)e^{-ikt} + w_1 + w_2 \tag{15}$$

is the solution of the problem (10) if $w_j$, $j = 1, 2$, are solutions of the following problems

$$\begin{cases} \partial_t^2 w_j - \Delta w_j = f & e^{-ikt},\ x \in \Omega,\ t \in I\!R; \\ \frac{\partial w_j}{\partial n}\big|_\Gamma = 0, & t \in I\!R;\ \ w_j = 0,\ t < 0 \end{cases} \tag{16}$$

where

$$f_1 = f_1(x, k) = (\Delta + k^2)\psi \tag{17}$$

$$f_2 = f_2(t, x, k) = (\beta - 1)f_1 - \psi(\beta'' - 2ik\beta') \tag{18}$$

From (17) and (18) it follows that

$$f_1(x, k) = 0 \text{ when } |x| > a, \text{ and } \|f_1\|_{L_2(\Omega)} \le C(1 + k^2) \tag{19}$$

(20) as well as $f_2(t, x, k) = 0$ when $|x| > a$ or $t > 1$, and $\|f_2\|_{L^2(\Omega)} \le C(1 + k^2)$.

It follows from (16), (20), Theorem 2 and Duhamel principle that

$$|\partial_t^j \partial_x^\alpha w_2| \le C(1 + k^2)|\partial_t^j \gamma(t)|,\ |x| \le a,\ t \ge T_0 + 1$$

where $C = C(h, a, j, \alpha)$ does not depend on $k$. It evidently follows from here that for $w_2$ the estimate (14) is valid with any $T \ge T_0 + 1$.

Further, let $w$ be the solution of the problem (11) with $f = f_1$. Then from the Duhamel principle it follows that

$$w_1 = e^{-ikt}\int_0^t w(\tau,x)e^{ik\tau}d\tau = e^{-ikt}\lim_{k_1\to k+i0}\int_0^t w(\tau,x)e^{ik_1\tau}d\tau. \tag{21}$$

If $\operatorname{Im} k_1 > 0$ the function

$$u(x,k_1) = \int_0^\infty w(\tau,x)e^{ik_1\tau}d\tau$$

belongs to $L^2(\Omega)$ and is the solution of the problem

$$(\Delta + k_1^2)u = -f_1,\ x\in\Omega;\ \frac{\partial u}{\partial n}|_\Gamma = 0.$$

We can now invoke the principle of limiting absorption which implies that

$$\lim_{k_1\to k+i0}\int_0^\infty w(\tau,x)e^{ik_1\tau}d\tau = u_1(x,k)$$

where $u_1$ is the solution of the problem

$$(\Delta + k^2)u_1 = -f_1,\ x\in\Omega;\ \frac{\partial u}{\partial n}|_\Gamma = 0;\ \frac{\partial u}{\partial n} - iku = o(r^{\frac{1-n}{2}}),\ r\to\infty. \tag{22}$$

This means that we can rewrite formula (21) in the following way,

$$\begin{aligned} w_1 &= e^{-ikt}u_1 - e^{-ikt}\lim_{k\to k_1+i0}\int_t^\infty w(\tau,x)e^{ik_1\tau}d\tau = \\ &= e^{-ikt}u_1 - \sum_{s=0}^{N+1}\frac{w_t^{(s)}(t,x)}{(-ik)^{s+1}} + e^{-ikt}\int_t^\infty \frac{w^{(N+2)}(\tau,x)}{(-ik)^{N+2}}e^{ik\tau}d\tau. \end{aligned} \tag{23}$$

where $|x|\le a,\ t > T_0 = T_0(a)$. For the second equality when we integrate by parts, we used the following estimate which is a consequence of (19) and Theorem 2

$$|\partial_t^j\partial_x^\alpha w| \le C(1+k^2)|\partial_t^j\gamma(t)|,\ |x|\le a,\ t\ge T_0. \tag{24}$$

From (23) and (24) it follows that for $w_1 - e^{-ikt}u_1$ the estimate (14) is valid for any $T \geq T_0$. As we proved estimate (14) for $w_2$ we have that

$$v = e^{-ikt}(\psi + u_1) + v_1$$

where $v_1$ satisfies estimate (14). It remains to note that from (17) and (22), it follows that function $u = \psi + u_1$ is the solution of the problem (1). Thus Theorem 3 is proved.

We will need the following simple lemma to prove Theorem 1. Recall that $S(x)$ is the distance between $x$ and the boundary $\Gamma$ and let

$$v_N = e^{ik(S(x)-t)} \sum_0^N a_j(t,x)(ik)^{-j-1} \tag{25}$$

where $a_j$ are infinitely smooth functions of $x \in \bar{\Omega}$ and $t \in I\!R$.

**Lemma 1.** There exist functions $a_j \in C^\infty$, $x \in \bar{\Omega}, t \in I\!R$, such that for any $N$ the following assertions are valid:

1. $(\partial_t^2 - \Delta)v_N = k^{-N-1}e^{ik(S(x)-t)}b_N(t,x)$, where $b_N$ is an infinitely smooth function,
2. $\frac{\partial v_N}{\partial n}\big|_\Gamma = e^{-ikt}(\beta(t)h + (ik)^{-N-1}\frac{\partial a_N}{\partial n}\big|_\Gamma)$,
3. $v_N = 0$ when $t \leq 0$,
4. The functions $a_j$ are independent of $t$ when $t \geq |x| + a + 1$.

**Proof.** This lemma is the outcome of standard WKB method [7], [4]. The Hamilton-Jacobi equation which corresponds to the operator $\partial_t^2 - \Delta$ is

$$\left(\frac{\partial \phi}{\partial t}\right)^2 = |\nabla \phi|^2.$$

The function $\phi = S(x) - t$ satisfies this equation. Hence the first assertion of the lemma will be satisfied if functions $a_j$ satisfy the transport equations. Let us write them. Geometro-optic rays $\ell$ which correspond to the phase function $\phi$ are straight rays which are orthogonal to $\Gamma$. Let

$$x = x_0 + \hat{\mathbf{n}}s,\ x_0 \in \Gamma,\ s \geq 0$$

be equations of these rays. Here $\hat{\mathbf{n}}$ is the unit vector of the exterior normal. It is obvious that $s = S(x)$. The transport equations are the following

$$(26) \qquad \begin{cases} \frac{\partial a_0}{\partial t} + \frac{\partial a_0}{\partial n} + \frac{1}{2}(\Delta S)a_0 = 0 \\ \frac{\partial a_j}{\partial t} + \frac{\partial a_j}{\partial n} + \frac{1}{2}(\Delta S)a_j = \frac{1}{2}\left(\frac{\partial^2 a_{j-1}}{\partial t^2} - \Delta a_{j-1}\right),\ j > 0. \end{cases}$$

Here

$$\frac{\partial a}{\partial n} = (\hat{\mathbf{n}}, \nabla a),\ x \in \ell.$$

The equations (26) are linear ordinary differential equations along space-time rays $\tilde{\ell} \subset I\!R^{n+1}$, $\tilde{\ell} = \tilde{\ell}(t_0, x_0) = \{(t, x) :\ x = x_0 + \hat{\mathbf{n}}s,\ t = t_0 + s,\ s \geq 0\},\ x_0 \in \Gamma,\ t_0 \in I\!R$.

So we define the functions $a_j$ as the solutions of the equations (25) which satisfy the following initial conditions

$$a_0(t_0, x_0) = \beta(t_0)h(x_0),\ a_j(t_0, x_0) = -\frac{\partial a_{j-1}}{\partial n}(t_0, x_0),\ j > 0.$$

It is now very easy to check that not only the first but also the other assertions of the Lemma are valid. This completes the proof.

**Proof of Theorem 1.** From Lemma 1 and well-known estimates of solutions of the mixed problem for wave equation [8] it follows that the solution of the problem (10) has the following form

$$(27) \qquad v = e^{ik(S(x)-t)}[\sum_0^N a_j(t, x)(ik)^{-j-1} + \tilde{a}_N(t, x, k)]$$

where $N$ is arbitrary and

$$|\partial_x^\alpha \tilde{a}_N| \leq Ck^{-N-2},\ k \geq 1,\ |x| \leq a,\ t \leq T+1 \tag{28}$$

for any $a, T$ and some constant $C = C(\alpha, a, T)$.

How we fix $a$ and choose $T$ such that

$$T \geq \max(T_0 + 1,\ 2a + 1)$$

where $T_0$ is the constant which is defined in Theorem 2. From Theorem 3 it follows that

$$|\partial_x^\alpha \int_T^{T+1} (u - e^{ikt}v)\varphi(t-T)dt| \leq Ck^{-N},\ |x| \leq a,\ k \geq 1. \tag{29}$$

Theorem 1 follows from (27) - (29) because the functions $a_j(x)$ are independent of $t$ when $|x| < a$, and $t > T$.

We now consider the asymptotic behavior of the far field coefficient of the solution in the case that the boundary value is independent of the parameter $k$.

**Theorem 4.** If function $h$ does not depend on $k$, then function (2) has the following asymptotic behavior as $k \to \infty$

$$f(\theta, k) = \frac{1}{ik} h(P^{-1}\theta)\kappa^{-1/2}(P^{-1}\theta)e^{-i<P^{-1}\theta,\ \theta>k}(1 + 0(k^{-1})).$$

**Remark.** It is not difficult to write full asymptotic expansion of the function $f$.

**Proof.** The function $f$ is defined by (2) in which the function $Th$ is that given in the corollary to Theorem 1. Asymptotic behavior of integrals of the type (2) as

$k \to \infty$ can be obtained with the help of stationary phase method. In particular Theorem 4 will be obtained if we apply Theorem 9 of Chapter 4 from [4] to the integral (2). Theorem 4 follows from these remarks.

## III. Proof of the Main Theorem

From Green's formula it follows that

$$\int\limits_{S^{n-1}} |f|^2 d\theta = -k^{-1} \operatorname{Im} \int\limits_{\Gamma} u \frac{\partial \bar{u}}{\partial n} dS.$$

Hence

$$\int\limits_{S^{n-1}} |f|^2 d\theta \leq \frac{1}{2}[\int\limits_{\Gamma} |u|^2 dS + \int\limits_{\Gamma} |k^{-1} \frac{\partial u}{\partial n}|^2 dS] = \frac{1}{2} |||h|||^2.$$

It is obvious that relations (5) follow from here.

Now let the function $h$ from boundary value problem (1) be equal to $k^{-1} h_\epsilon$ where $h_\epsilon$ is defined in (6). Then $h$ is independent of $k$ and Theorems 1 and 3 are valid for this function $h$. In particular from the corollary to Theorem 1 it follows that

$$Th_\epsilon(x) = kTh(x) = -ih(x)(1 + 0(k^{-1})) = (ik)^{-1} h_\epsilon(x)(1 + 0(k^{-1})).$$

Hence

$$|||h_\epsilon||| = \frac{\sqrt{2}}{k} \|h_\epsilon\|_{L^2(\Gamma)} (1 + 0(k^{-1})) = (\int\limits_{\Gamma} |g_\epsilon(Px)|^2 \kappa(x) dS)^{1/2} (1 + 0(k^{-1})).$$

As the Jacobian of the mapping $P$ is equal to $\kappa^{-1}(x)$ the last equality can be rewritten in the form

$$|||h_\epsilon||| = (\int\limits_{S^{n-1}} |g_\epsilon(\theta)|^2 dS)^{1/2} (1 + 0(k^{-1})).$$

This and the first of the relations (4) lead to (7).

As we already mentioned Theorem 4 is valid if $h = k^{-1}h_\epsilon$. If we multiply boundary function $h$ by $k$ then function $f$ is multiplied by $k$ as well. Hence Theorem 4 is valid if $h = h_\epsilon$ and therefore

$$\int\limits_{S^{n-1}} |f|^2 a(\theta)d\theta = \frac{1}{k}\int\limits_{S^{n-1}} |h_\epsilon(P^{-1}\theta)|^2\kappa^{-1}(P^{-1}\theta)a(\theta)d\theta(1+0(k^{-1})).$$

From here and (6) we obtain

$$\int\limits_{S^{n-1}} |f|^2 a(\theta)d\theta = \frac{1}{2}\int\limits_{S^{n-1}} |g_\epsilon(\theta)|^2 a(\theta)d\theta(1+0(k^{-1})).$$

This and the second of the relations (4) lead to (8).

Thus, the main theorem is proved.

References


1. T. S. Angell and R. E. Kleinman, Optimal control problems in radiation and scattering, in *Symposium of Applied Mathematics dedicated to the late Professor Dr. R. Timman*, A. J. Hermans, M. W. C. Oosterveld, editors, Delft University Press, Sijthoff & Noordhoff International Publishers, Delft, 1978, 78–90.

2. T. S. Angell and R. E. Kleinman, Generalized exterior boundary value problems and optimization for the Helmholtz equation, *J. Optimization Theory Appl.*, **37** (1982), 469–497.

3. B. R. Vainberg, Quasi–classical approximation in stationary scattering problems, *Funct. Anal. Appl.*, **11**, pp. 247–257, (1977).

4. B. R. Vainberg, *Asymptotic Methods in Equations of Mathematical Physics*, Moscow State Univ., 1982; English Translation: Gordon and Breach Publishers, 1989.

5. A. Majda, High frequency asymptotics for the scattering matrix and the inverse problem of acoustical scattering, *Comm. Pure Appl. Math.*, **29**, pp. 261–291, (1976).

6. A. Majda and M. Taylor, The asymptotic behavior of the diffraction peak in classical scattering, *Comm. Pure Appl. Math.*, **30**, pp. 639–669, (1977).

7. V. Guillemin and S. Sternberg, Geometric asymptotic, Providence, RI, Amer. Math. Soc., 1977.

8. M. Taylor, Pseudodifferential operators, Princeton University Press, Princeton, 1981.

9. B. Vainberg and V. Grushis, Uniformly nonelliptic problems, *Math. USSR - Sbornik*, **2** N1, pp. 111–133, (1967).

10. G. Ulhmann, Inverse boundary value problems and applications, *Asterisque*, **207**, pp. 153–211, (1992).

# "BLIND" SHAPE RECONSTRUCTION FROM EXPERIMENTAL DATA

**P. M. van den Berg**
Laboratory of Electromagnetic Research
Faculty of Electrical Engineering
Delft University of Technology
Delft, The Netherlands

**M. B. Coté**
Rome Laboratory
ERCT
Hanscom AFB, MA 01731–5000

and

**R. E. Kleinman**
Center for the Mathematics of Waves
Department of Mathematical Sciences
University of Delaware
Newark, DE 19716

**Technical Report No. 94–2**

Submitted to: IEEE-AP Trans. Antennas and Prop.

# "Blind" Shape Reconstruction from Experimental Data

**P.M. van den Berg**

*Laboratory of Electromagnetic Research, Faculty of Electrical Engineering*
*Delft University of Technology, Delft, The Netherlands*

**M.G. Coté**

*Rome Laboratory/ERCT,*
*Hanscom AFB. MA 01731-5000. U.S.A.*

R.E. Kleinman

*Center for the Mathematics of Waves, Department of Mathematical Sciences*
*University of Delaware, Newark, DE 19716, U.S.A.*

A method for reconstructing the shape of a bounded impenetrable object from measured scattered field data is presented. The reconstruction algorithm is in principle the same as that used before for reconstructing the conductivity of a penetrable object and uses the fact that for high conductivity the skin depth of the scatterer is small, in which case the only meaningful information produced by the algorithm is the boundary of the scatterer. A striking increase in efficiency is achieved by incorporating into the algorithm the fact that for large conductivity the contrast is dominated by a large positive imaginary part. This fact together with the knowledge that the scatterer is constrained in some test domain constitute the only a priori information about the scatterer that is used. There are no other implicit assumptions about the location, connectivity, convexity or boundary conditions. The method is shown to successfully reconstruct the shape of an object from experimental scattered field data in a "blind" test.

## I. INTRODUCTION

The present paper describes a successful example of the reconstruction of the shape of a scattering object from experimentally determined scattering data. In contrast with other inversion methods, the reconstruction is accomplished from real rather than synthetic data, so there is no chance of even inadvertently committing the "inverse crime" of using the same numerical method in the inversion algorithm as is used for solving the forward or direct problem to produce the synthetic "measured" data. The possibility of favorably prejudicing the outcome of the inversion algorithm was eliminated by a "blind" use of the measured data in the inversion algorithm; that is, knowledge of the geometry of the object from which the scattered field was measured was not supplied to those running the algorithm until after the reconstruction was completed.

The reconstruction algorithm is that described by Kleinman and Van den Berg [1], in which an iterative algorithm for the reconstruction of complex contrast profiles [2, 3] is adapted to reconstructing the shape and location of a perfectly conducting scatterer by making the assumption that the unknown contrast is essentially non-negative imaginary. The experimental data were obtained on the Ipswich Test Range of Rome Laboratories [4].

## II. DESCRIPTION OF THE METHOD

Assume that a two-dimensional conducting obstacle $D$ is irradiated successively by a number ($j = 1, \cdots, J$) of known incident fields with the electric-field vector parallel to the cylindrical object (TM-case). For each excitation, we then have a scalar problem and the incident electric-field component is denoted as $u_j^{inc}$ and the total electric-field component is denoted as $u_j$. For each excitation, the direct scattering problem may be reformulated

as the domain integral equation

$$L_{(\chi)}u_j(\boldsymbol{p}) := u_j(\boldsymbol{p}) - G_D\chi u_j(\boldsymbol{p}) = u_j^{inc}, \quad \boldsymbol{p} \in D\,, \tag{1}$$

where

$$G_D\chi u_j(\boldsymbol{p}) := k^2 \int_D G(\boldsymbol{p},\boldsymbol{q})\chi(\boldsymbol{q})u_j(\boldsymbol{q})dv_{\boldsymbol{q}}, \quad \boldsymbol{p} \in D\,, \tag{2}$$

and

$$G(\boldsymbol{p},\boldsymbol{q}) = \frac{i}{4}H_0^{(1)}(k|\boldsymbol{p}-\boldsymbol{q}|)\,. \tag{3}$$

Here, $k$ is the wavenumber, $\chi$ is taken to be equal to $i\zeta^2$ for real $\zeta$ ($\zeta^2 = \sigma/\omega\varepsilon_0$), and $\boldsymbol{p}$ and $\boldsymbol{q}$ are two-dimensional position vectors. $G(\boldsymbol{p},\boldsymbol{q})$ is the free-space Green's function in two dimensions. $G_D$ is an operator mapping $L^2(D)$ (square integrable functions in $D$) into itself. If $S$ is a surface enclosing $D$ then the scattered electric-field component $u_j^{sct} = u_j - u_j^{inc}$ on $S$, is given by $G_S\chi u_j$ where $G_S$ is the same operator defined in (2), except the field point $\boldsymbol{p}$ now lies on $S$. Hence $G_S$ is an operator mapping $L^2(D)$ into $L^2(S)$. We assume that $u_j^{sct}$ is measured on $S$ and denote by $f_j(\boldsymbol{p})$, $\boldsymbol{p} \in S$, the measured data for each excitation $j$, $j = 1, \cdots, J$. The conductivity reconstruction problem is that of finding $\chi$ for given $f_j$, or solving the equations

$$G_S\chi u_j(\boldsymbol{p}) = f_j(\boldsymbol{p}), \quad \boldsymbol{p} \in S, \quad j = 1, \cdots, J\,, \tag{4}$$

for $\chi$, subject to the additional condition that $u_j$ and $\chi = i\zeta^2$ satisfy (1) in $D$ for each $j$.

Specifically we use the iterative construction of sequences $\{u_{j,n}\}$ and $\{\zeta_n\}$ as follows:

$$u_{j,n} = u_{j,n-1} + \alpha_n v_{j,n}, \quad \zeta_n = \zeta_{n-1} + \beta_n\xi_n, \quad n = 1, 2, \cdots\,. \tag{5}$$

For each $n$, the functions $v_{j,n}$ and $\xi_n$ are update directions for the functions $u_{j,n}$ and $\zeta_n$, respectively, while the complex parameter $\alpha_n$ and the real parameter $\beta_n$ are weights to be determined. The residual errors at each step in the state equation and data equation are defined as

$$r_{j,n} = u_j^{inc} - L_{(i\zeta_n^2)}u_{j,n}, \quad \rho_{j,n} = f_j - iG_S\zeta_n^2 u_{j,n}\,, \tag{6}$$

The constants $\alpha_n$ and $\beta_n$ are determined by minimizing the value of the cost functional

$$F_n = w_D \sum_{j=1}^{J} \|r_{j,n}\|_D^2 + w_S \sum_{j=1}^{J} \|\rho_{j,n}\|_S^2 , \tag{7}$$

in which

$$w_D = \left( \sum_{j=1}^{J} \|u_j^{inc}\|_D^2 \right)^{-1} \quad \text{and} \quad w_S = \left( \sum_{j=1}^{J} \|f_j\|_S^2 \right)^{-1} . \tag{8}$$

where the subscripts $S$ and $D$ are included in the norm $\|\cdot\|$ in $L^2$ to indicate the domain of integration. Substitution of Eqs. (5)-(6) in the cost functional of Eq. (7) results in an expression involving terms determined at the $(n-1)$-st step, the directions $\xi_n$ and $v_{j,n}$, and the two parameters $\alpha_n$ and $\beta_n$. Once the directions $\xi_n$ and $v_{j,n}$ are chosen, we have a nonlinear expression in $\alpha_n$ and $\beta_n$. The values of the parameters $\alpha_n$ and $\beta_n$ are determined by requiring $F_n$ to be a minimum. Minimization of the quantity $F_n$ is accomplished by solving this non-linear problem in $\alpha_n$ and $\beta_n$ using a standard conjugate gradient method.

The update directions $v_n$ and $\xi_n$ are chosen as the Polak-Ribière conjugate gradient directions as in [1], namely

$$v_{j,n} = g_{j,n}^v + \gamma_n^v v_{j,n-1} \quad \text{and} \quad \xi_n = g_n^\xi + \gamma_n^\xi \xi_{n-1} . \tag{9}$$

where

$$\gamma_n^v = \frac{\sum_{j=1}^{J} \langle g_{j,n}^v , g_{j,n}^v - g_{j,n-1}^v \rangle_D}{\sum_{j=1}^{J} \|g_{j,n-1}^v\|_D^2} \quad \text{and} \quad \gamma_n^\xi = \frac{\langle g_n^\xi , g_n^\xi - g_{n-1}^\xi \rangle_D}{\|g_{n-1}^\xi\|_D^2} , \tag{10}$$

and the gradients are given by

$$g_{j,n}^v = w_D(r_{j,n-1} + i\zeta_{n-1}^2 \overline{G}_D r_{j,n-1}) - i w_S \zeta_{n-1}^2 \overline{G}_S \rho_{j,n-1} \tag{11}$$

and

$$g_n^\xi = 2\zeta_{n-1} \mathrm{Im} \left[ w_D \sum_{j=1}^{J} \overline{u}_{j,n-1} \overline{G}_D r_{j,n-1} - w_S \sum_{j=1}^{J} \overline{u}_{j,n-1} \overline{G}_S \rho_{j,n-1} \right] . \tag{12}$$

The operators $\overline{G}_D$ and $\overline{G}_S$ are the adjoints of $G_D$ and $G_S$, respectively, mapping $L^2(D)$ and $L^2(S)$ into $L^2(D)$.

The initial estimates $u_{j,0}$ and $\zeta_0$ are chosen as in [1] to be

$$u_{j,0} = u_j^{inc} + G_D w_{j,0} , \tag{13}$$

where

$$w_{j,0} = \frac{\sum_{k=1}^{J} < f_k, G_S \overline{G}_S f_k >_S}{\sum_{k=1}^{J} \| G_S \overline{G}_S f_k \|_S^2} \overline{G}_S f_j , \tag{14}$$

and

$$\zeta_0^2 = \left( \frac{\sum_{j=1}^{J} \frac{\{\mathrm{Im}[w_{j,0} \overline{u}_{j,0}]\}^2}{|u_{j,0}|^2}}{\sum_{j=1}^{J} |u_{j,0}|^2} \right)^{\frac{1}{2}} . \tag{15}$$

## III. EXPERIMENTAL SETUP

Here we describe how the field scattered by the mystery object was measured and calibrated for the reconstruction. The measurement frequency was 10.0 GHz, thus the wavelength ($\lambda$) was 3 cm. Bistatic scattering measurements were made in a plane perpendicular to the axis of the cylindrical object, 30 cm (10 $\lambda$) in length and the measurement plane intersected at mid length. For convenience, a Cartesian coordinate system was oriented with $z$ along the cylinder axis, and measurements were made in the $(x, y)$ plane. The measurement configuration is shown in Fig. 1. The scattered fields were collected for incident angles of $\phi^i = \{0, 5, 10, 15, 20, 45, 60, 90\}$ degrees, over the observation sector $0 \leq \phi^s \leq 359.5°$ with a sample spacing, $\Delta\phi^s = 0.5°$.

The object and transmit antenna were fixed for each $\phi^i$ and the receive antenna was rotated on a semi-circular arc about the object from back scatter to forward scatter recording the total field coincident with the receive antenna polarization. A second measurement was made with the object removed. This background field measurement was subtracted

from each of the total-field measurements to obtain measured data proportional to the scattered field. The range from the transmit antenna to the object was 3.7 m and the range from the object to the receive antenna aperture was 2.8 m. Both the source and the probe antennas had circular apertures 15.24 cm in diameter. With these measurement ranges and antennas, the object illumination was uniform in magnitude to within 0.2 dB along the $x$ direction and 1 dB along the $z$ direction. The illumination phase taper over the object was approximately 10° and 50° in the $x$ and $z$ directions, respectively. We note that both the end sides ($z = \pm 15$ cm) of the finite cylindrical object were illuminated quite strongly and therefore one might expect the measured scattered field would contain an undesirable diffraction from the edges of the end sides. However in this experiment the planes of incidence and observation were always normal to the $z$ axis, which insured that the scattered field was dominated by the specular response and the diffraction from the two truncating sides, being much less, was not observable. Thus, the measured scattering from the finite cylindrical object was very close to that from an infinite cylindrical object.

The measurement system used can only scan over a 190° bistatic angular sector. This means that in order to get scattering data over a complete 360° bistatic observation sector, two measurement runs had to be made for each incident direction, one measurement run to cover the observation sector, $\phi^i - 5° \leq \phi^s \leq \phi^i + 185°$, and the other to cover, $\phi^i + 175° \leq \phi^s \leq \phi^i + 365°$. The data from each measurement run must be independently calibrated and then spliced together to make a complete data set. In this experiment coverage of the first observation sector for every incident angle of interest (except $\phi^i = 0$), was accomplished by measurements made in March 1990 . The second observation sector was obtained for all incident angles of interest (except $\phi^i = 90°$), by measurements made in October 1991. The instrumentation radar used in the October 91 measurements was more sensitive than the radar used in the March 90 measurements so that the March 90 portion of each complete data set had an uncertainty significantly greater than the

October 91 portion. In addition to a variable uncertainty each data set contained a sector of completely erroneous scattering centered about the back-scattering direction ($\phi^i - 5° \leq \phi^s \leq \phi^i + 5°$) caused by the interruption of the object illumination when the receive antenna passed between the transmit antenna and object. For each measurement run we filled in the erroneous back-scattering region by extrapolating the complex data on $\phi^i + 5° \leq \phi^s \leq \phi^i + 185°$, using a least squares linear prediction algorithm [5].

The raw scattering data resulting from the phasor subtraction of the total-field and background measurements, has a magnitude proportional to the object scattering cross section per unit length, and a phase proportional to the phase of the scattered electric field referenced to the center of rotation of the bistatic positioner. Aligning the object so that its symmetry axis coincides with this rotation axis is practically impossible. Our calibration procedure must compensate for the phase error caused by this misalignment in addition to calibrating the magnitude. We calibrated the scattering from the object by the following procedure. We computed a point calibration phasor,

$$\Psi(\phi^s) = \frac{\pi}{2\lambda} \frac{P^{exp}(\phi^s)}{P^{comp}(\phi^s)}. \tag{16}$$

In (16), $P^{exp}(\phi^s)$ is the measured scattered pattern of the object for a particular measurement run, and $P^{comp}(\phi^s)$ is the far-field scattering pattern computed for an infinitely long cylinder that approximates the present object, but with its symmetry axis at the $z$ axis. From the calibration phasor we compute an average calibration factor, $\Psi_0$, and three constants, $a$, $b$, and $c$. The average calibration factor is defined as

$$\Psi_0 = \frac{1}{N} \sum_{n=1}^{N} \left| \Psi(n\Delta\phi^s + \phi^i + 5°) \right| , \tag{17}$$

where, $N$ is the number of data points in the given measurement run excluding the erroneous data in the 10° back-scattering sector. The three constants are determined such that they produce the best fit (in the least square sense) to the expression,

$$\arg\left[\Psi(\phi^s)\right] = a + b\cos(\phi^s + c) . \tag{18}$$

This curve fitting step is needed to correct for the misalignment phase error (see [4] for a more thorough discussion). With those four constants computed for each measurement run the calibrated scattering cross section per unit length, $\sigma^{cal}(\phi^s)$, was calculated from the relation,

$$\sigma^{cal}(\phi^s) = \Psi_0 P^{exp}(\phi^s) \exp\{-i[a + b\cos(\phi^s + c)]\} . \tag{19}$$

In this way we have arrived at experimental data that belongs to the object with the symmetry axis coinciding with the $z$ axis. We note that this procedure was necessary, because we have only angles of incidence in a quarter plane, and using the symmetry, we can obtain scattered data from angles of incidence in the full plane. These experimental data are recalibrated for use in the inversion algorithm as described in the next section.

## IV. RECONSTRUCTION

The measurement surface $S$ is chosen to be a circle containing the test domain. We assume that the radius of this circle is large enough so that the far-field approximation of (4) may be employed, and the far-field coefficient is the quantity of interest so that the dependence on the radius is removed. In that case the data may be written as

$$f_j(\boldsymbol{p}) \sim \frac{i}{4}\left(\frac{2}{\pi k|\boldsymbol{p}|}\right)^{\frac{1}{2}} \exp(ik|\boldsymbol{p}| - i\tfrac{\pi}{4})\, f_j^{\infty}(\hat{\boldsymbol{p}}) , \tag{20}$$

and the data equation (4) may be replaced by

$$\int_D \exp(-ik\hat{\boldsymbol{p}}\cdot\boldsymbol{q})\chi(\boldsymbol{q})u_j(\boldsymbol{q})dv_{\boldsymbol{q}} = f_j^{\infty}(\hat{\boldsymbol{p}}) , \quad \hat{\boldsymbol{p}} \in S . \tag{21}$$

where $\hat{\boldsymbol{p}}$ is the unit vector in the direction of observation and $S$ now denotes the space of these unit vectors, the unit circle. Further, $f_j^{\infty}(\hat{\boldsymbol{p}})$ are the measured far-field data. In the examples, we take from the measured far-field data the values at 36 angles equally spaced around the object (the domain $S$ consists of 36 discrete points $\hat{\boldsymbol{p}}_l$). In the experiments

only 8 excitations are carried out. The incident fields are approximated as plane waves incident at an angle of 0, 5, 10, 15, 20, 45, 60 and 90 degrees with the $x$-axis, respectively. To obtain scattered-field data from incident waves distributed around the object, we take advantage of the a priori information that the mystery object is symmetric with respect to the planes $x = 0$ and $y = 0$. Doing so we obtain scattered-field data from 28 excitations ($J = 28$).

We further have a priori information that the mystery object lies inside a circle with a radius of 0.060 m and the frequency of operation is 10 GHz. We therefore will assume that the object is located inside a test square divided into $63 \times 63$ subsquares of $0.002 \times 0.002$ $m^2$. The discretized version of the algorithm is discussed in [1].

*Calibration*

To test the computer code, we first run the algorithm for synthetic data obtained in the well-known problem of scattering of a plane wave by a perfectly conducting circular cylinder with origin at the center of the test square. We employ the same angles of incidence and data points as used in the experimental case. The analytic solution in terms of Bessel functions has been employed. The data are denoted as

$$f_j^{exl}(\hat{\boldsymbol{p}}_l) := f_j^{\infty}(\hat{\boldsymbol{p}}_l), \quad j = 1, \cdots, 28, \quad l = 1, \cdots, 36. \tag{22}$$

The radius, $a$, of this circular cylinder is 0.015 m. The wavelength is $\lambda = 0.030$ m, so that $ka = \pi$. We have seen that our scheme indeed reconstructs the location and the shape of a perfectly conducting cylinder by reconstructing the imaginary contrast at the boundary [1]. However, the reconstructed contrast at the boundary becomes highly oscillatory after a couple of iterations. The peaks appear to increase with the number of iterations and it becomes difficult to choose the level value of the contour that estimates the boundary of the object. The visualization of the boundary of the object is improved when we impose an upper bound to the reconstructed contrast. If at some point in the iteration

the reconstructed $\zeta_n$ is larger than $\zeta_{max}$, the contrast is replaced by $\zeta_{max}$. In our example we take $\zeta_{max} = 1$. Some surface plots of the reconstructed profiles (the imaginary part of the contrast, $\mathrm{Im}[\chi] = \zeta^2$) from the synthetic data of the circular cylinder are presented in Fig. 2. The result at 32 iterations has also been presented in Fig. 2a, where we have plotted the boundaries of the test domain and the contour lines $\zeta = 1$. The exact location of the boundary of the object is indicated by the dashed circle. The asymmetry of the choice of the incident angles of excitations is clearly visible in the reconstructed boundary. We observe that the boundary is located with an error of the sample width.

Next we measure experimentally the scattering from a circular cylinder with the same dimensions. These data are denoted as $f_j^{exp}(\hat{\boldsymbol{p}}_l)$, $j = 1, \cdots, 28$, $l = 1, \cdots, 36$. To calibrate an overall phase shift between the definition of the phase of the measurement data and the one defined in the reconstruction scheme (and to some extent the amplitudes), we assume that the measured signal is correct apart of a multiplicative complex factor and enforce the data to be

$$f_j^{cal}(\hat{\boldsymbol{p}}_l) := C\, f_j^{exp}(\hat{\boldsymbol{p}}_l), \quad j = 1, \cdots, 28, \quad l = 1, \cdots, 36\,. \tag{23}$$

The constant $C$ is determined from the analytical data pertaining to this object by minimizing the deviation

$$\sum_{j=1}^{J} \| f_j^{anl}(\hat{\boldsymbol{p}}) - C\, f_j^{exp}(\hat{\boldsymbol{p}}) \|_S^2 = \sum_{j=1}^{36} \sum_{l=1}^{28} | f_j^{anl}(\hat{\boldsymbol{p}}_l) - C\, f_j^{exp}(\hat{\boldsymbol{p}}_l) |^2\,, \tag{24}$$

resulting in

$$C = \frac{\sum_{j=1}^{J} \langle f_j^{anl}(\hat{\boldsymbol{p}}), f_j^{exp}(\hat{\boldsymbol{p}}) \rangle_S}{\sum_{j=1}^{J} \| f_j^{anl}(\hat{\boldsymbol{p}}) \|_S^2} = \frac{\sum_{j=1}^{36} \sum_{l=1}^{28} f_j^{anl}(\hat{\boldsymbol{p}}_l)\, \overline{f_j^{exp}(\hat{\boldsymbol{p}}_l)}}{\sum_{j=1}^{36} \sum_{l=1}^{28} | f_j^{exp}(\hat{\boldsymbol{p}}_l) |^2}\,, \tag{25}$$

where the overbar denotes complex conjugate. After substitution of the resulting numer-

ical value of $C$ into the deviation of Eq. (24), we found that

$$\left[\frac{\sum_{j=1}^{J}\|f_j^{cal}(\hat{\boldsymbol{p}}) - C\, f_j^{exp}(\hat{\boldsymbol{p}})\|_S^2}{\sum_{j=1}^{J}\|f_j^{cal}(\hat{\boldsymbol{p}})\|_S^2}\right]^{\frac{1}{2}} = 0.079\,, \tag{26}$$

that is, a mean square deviation of about 8 %. Using these recalibrated data, we run the inversion algorithm. Some surface plots of the reconstructed profiles (the imaginary part of the contrast, $\mathrm{Im}[\chi] = \zeta^2$) from these calibrated experimental data of the circular cylinder are presented in Fig. 3. The result at 32 iterations has also been presented in Fig. 3a, where we have plotted the boundaries of the test domain and the contour lines $\zeta = 1$. The reconstruction from our experimental data is not very different from the reconstruction using the synthetic data.

*Mystery object*

Observing that our reconstruction of the circular cylinder was successful, we now continue to reconstruct a mystery object from experimental data. The experimental data from this mystery object were first multiplied with the complex constant $C$, computed by minimizing the global deviation between analytical and experimental data from the circular-cylinder case. This ensures that an overall phase shift between the one defined in the measurements and the one in the reconstruction scheme is corrected. We then run the inversion algorithm and the results of the reconstruction are shown in Figs. 4 and 4a. It clearly shows that the mystery object is probably a strip of about a width of 12 cm and a thickness of less than or equal to 4 mm.

Finally, we show in Figs. 5 and 5a, the reconstruction in a larger test domain, divided into $63 \times 63$ subsquares of $0.004 \times 0.004$ m$^2$. The result of the reconstruction, using this coarser grid, is consistent with the previous result.

After this reconstruction, the mystery was revealed to those running the reconstruction

algorithm: the object is a 10 $\lambda$ long (30 cm), 4 $\lambda$ (12 cm) wide and 0.106 $\lambda$ (0.32 cm) thick aluminum plate. Obviously, the cross-sectional dimensions of the mystery object that are obtained from the reconstruction results are very close to the real ones.

In order to show the quality of the measurements, we have computed the far-field data of the infinitely long and infinitely thin strip using the eigenfunction expansions described by Asvestas and Kleinman [6]. In Fig. 6 we compare the computed results of the strip with the measured results of the plate for one incidence direction ($\phi^i = 10°$). Notice that the measured scattering from about 5° off back scatter ($\phi^s = 15°$) to about $\phi^s = 190°$ is noisier than the remainder of the curve. The noisy sector corresponds to the measurements made in March 1990. In addition notice that the measured curve near $\phi^s = \phi^i = 10°$, is flat and does not match the exact curve. This is the back-scattering region that contains the extrapolated data.

## V. CONCLUSIONS

This paper presents definitive evidence of the effectiveness of the modified gradient inverse scattering algorithm in reconstructing the shape of a perfectly conducting cylindrical object of arbitrary cross section from scattered field data. In earlier papers it was shown that the algorithm was effective in reconstructing the contrast of penetrable objects, the boundary of impenetrable circular cylinders, and was stable with respect to white noise. All previous tests were performed with synthetic, i.e. computer simulated, scattering experiments and thus were not free from the possibility that they were tainted by an "inverse crime" of somehow using knowledge of the scatterer to favorably influence the reconstruction. The present results show conclusively that the algorithm will yield a successful reconstruction when the data are obtained experimentally and the shape of the object was not known before the reconstruction was completed, thus removing any

question that an "inverse crime", however inadvertent, was committed. These results describe only one scattering experiment and additional experiments are needed, not only to reconfirm the present results, but also to test the effectiveness of the reconstruction algorithm for penetrable scatterers.

*Acknowledgments.* This work was supported under AFOSR Grant-91-0277 and NATO Grant-0230/88.

## References

[1] R.E. Kleinman and P.M. van den Berg, "Two-dimensional location and shape reconstruction," *Radio Science*, to be published, 1994.

[2] R.E. Kleinman and P.M. van den Berg, "A modified gradient method for two-dimensional problems in tomography," *Journal of Computational and Applied Mathematics*, vol. 42, pp. 17–35, 1992.

[3] R.E. Kleinman and P.M. van den Berg, "An extended range modified gradient technique for profile inversion," *Radio Science*, vol. 28, pp. 877–884, September-October 1993.

[4] M. G. Coté, "Automated swept-angle bistatic scattering measurements using continuous wave radar," *IEEE Trans. Instrum. Meas.*, vol. 41, pp. 185-192, April 1992.

[5] S. L. Marple Jr., *Digital Spectral Analysis*, Englewood Cliffs, NJ: Prentice-Hall, 1987, Ch. 8 [subroutine COVAR].

[6] J. S. Asvestas, and R. E. Kleinman, "The strip," Ch. 4 in *Electromagnetic and Acoustic Scattering by Simple Shapes*, J.J. Bowman, T.B.A. Senior, P.L.E. Uslenghi, Eds. Amsterdam: North-Holland, 1969.

## CAPTIONS OF FIGURES

Fig. 1. Schematic diagram of the automated swept-angle bistatic measurement system.

Fig. 2. The reconstructed imaginary values of the contrast from synthetic data.

Fig. 2a. Comparison between the reconstructed boundary and the exact one (synthetic data, $n = 32$).

Fig. 3. The reconstructed imaginary values of the contrast from experimental data.

Fig. 3a. Comparison between the reconstructed boundary and the exact one (experimental data, $n = 32$).

Fig. 4. The reconstructed imaginary values of the contrast of the mystery object (dimension of test domain = $0.126 \times 0.126$ m$^2$).

Fig. 4a. The reconstructed boundary of the mystery object ($n = 32$, dimension of test domain = $0.126 \times 0.126$ m$^2$.

Fig. 5. The reconstructed imaginary values of the contrast of the mystery object (dimension of test domain = $0.252 \times 0.252$ m$^2$).

Fig. 5a. The reconstructed boundary of the mystery object ($n = 64$, dimension of test domain = $0.252 \times 0.252$ m$^2$).

Fig. 6. Bistatic scattering from the $4\lambda$ strip illuminated 10° off grazing.

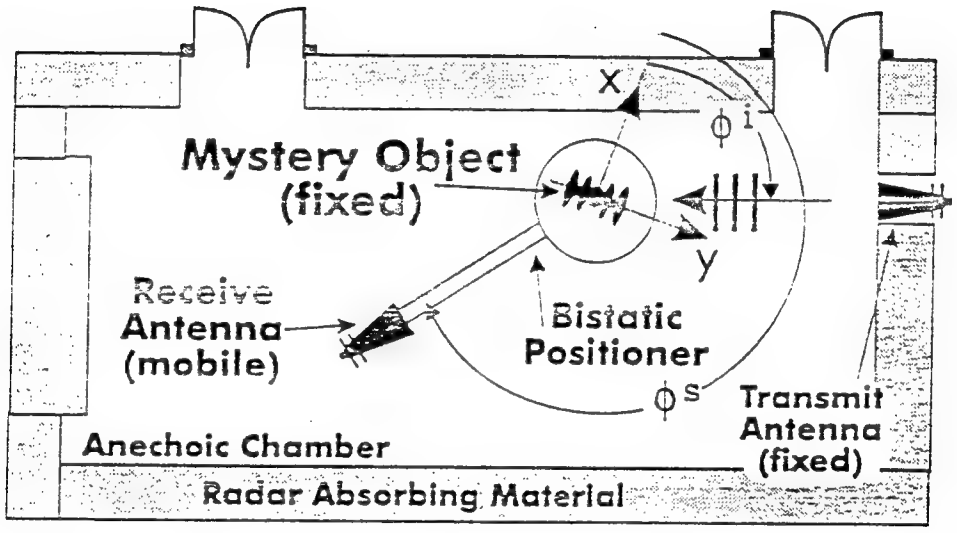

x
$\phi^i$
Mystery Object
(fixed)
y
Receive
Antenna
(mobile)
Bistatic
Positioner
$\phi^s$
Transmit
Antenna
(fixed)
Anechoic Chamber
Radar Absorbing Material

Figure 1

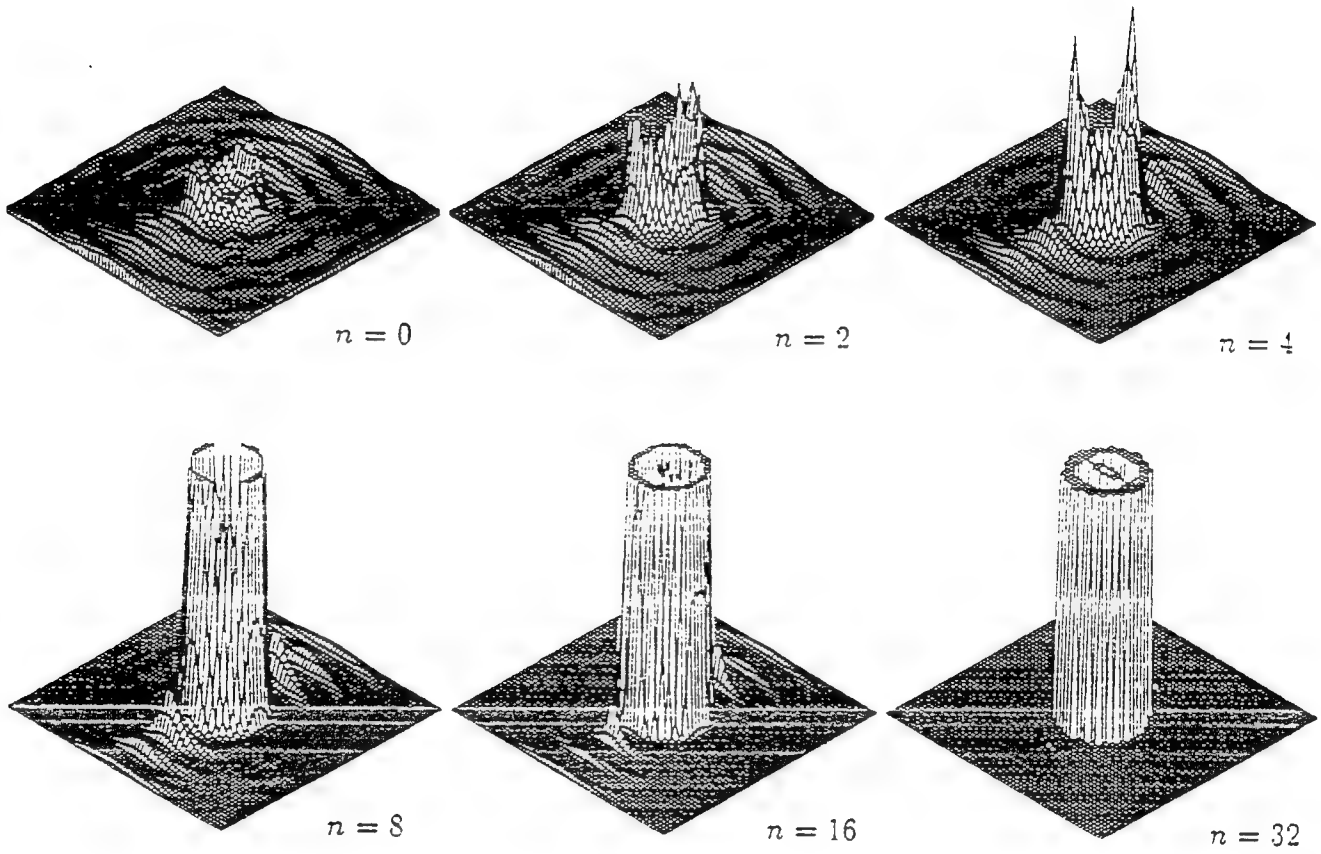



Figure 2

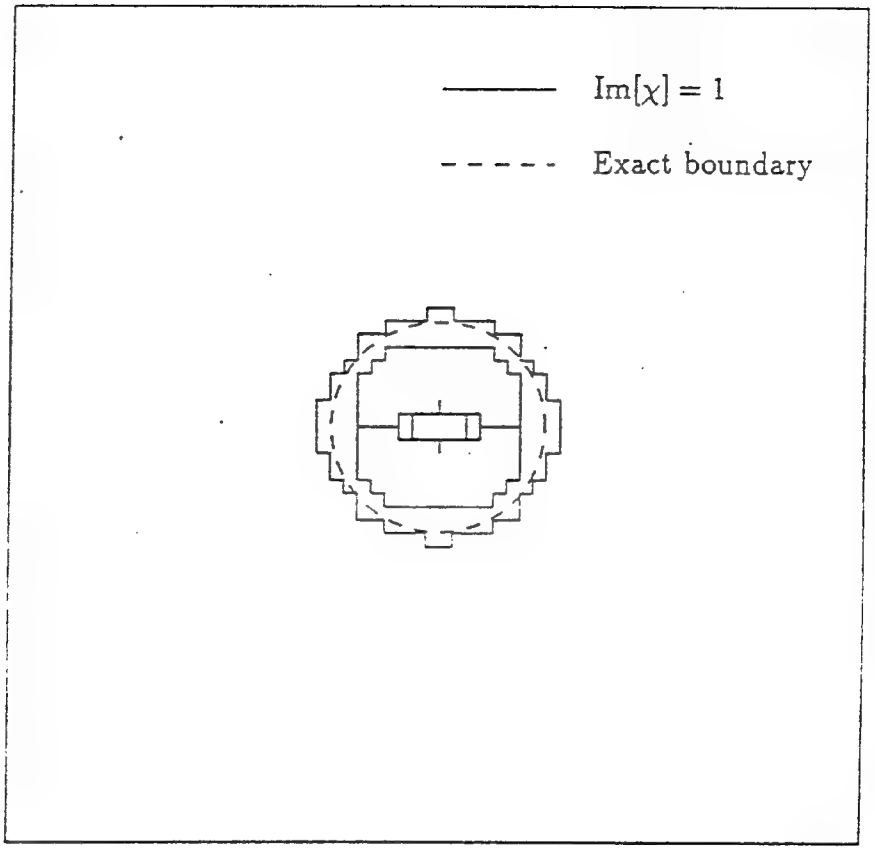



Figure 2a

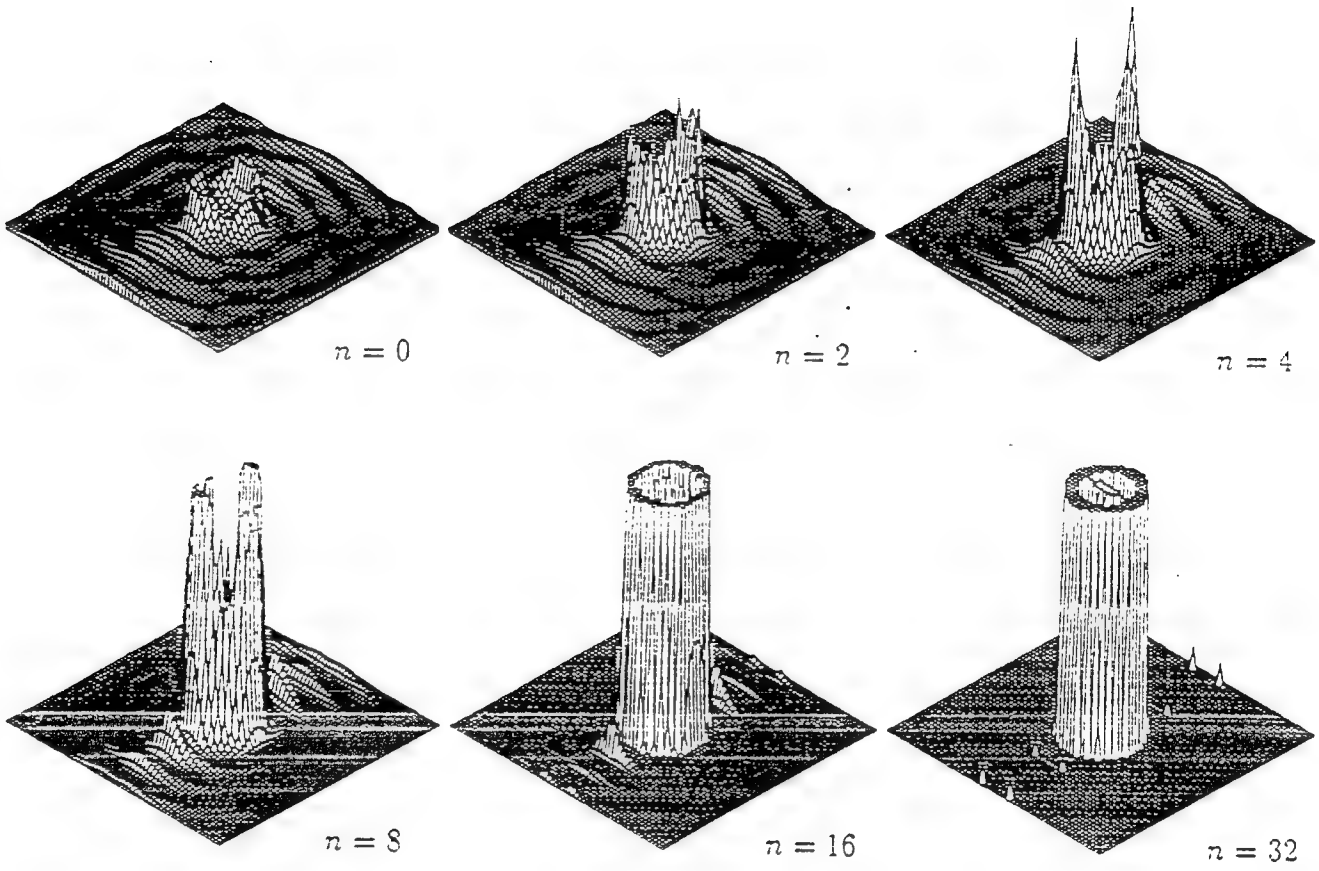



Figure 3

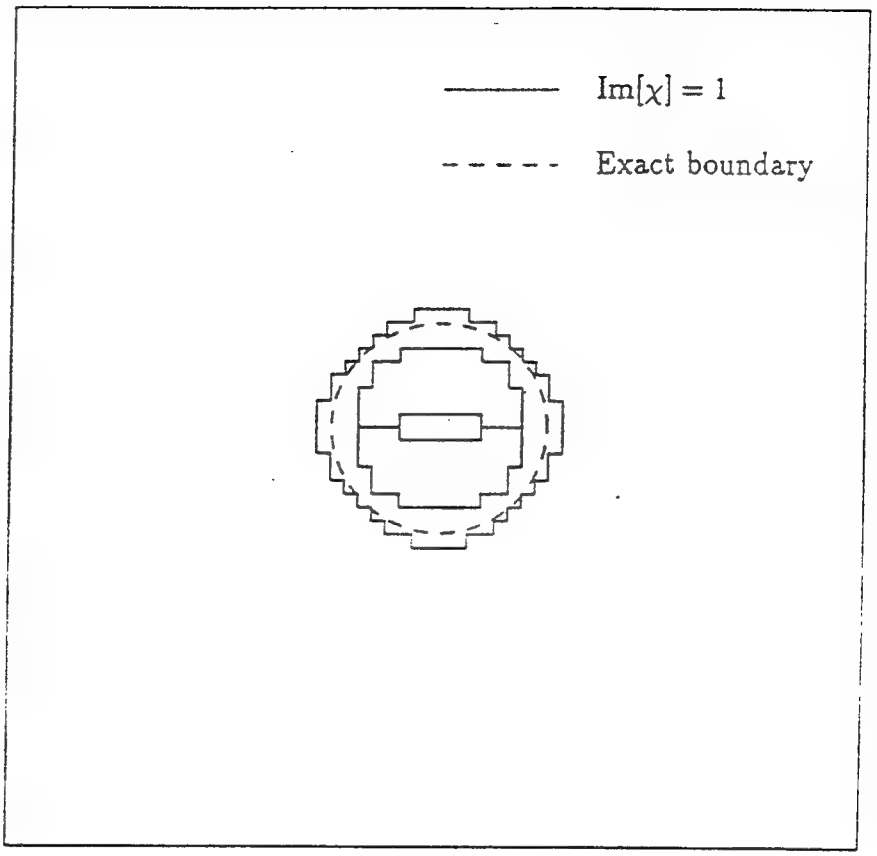



Figure 3a

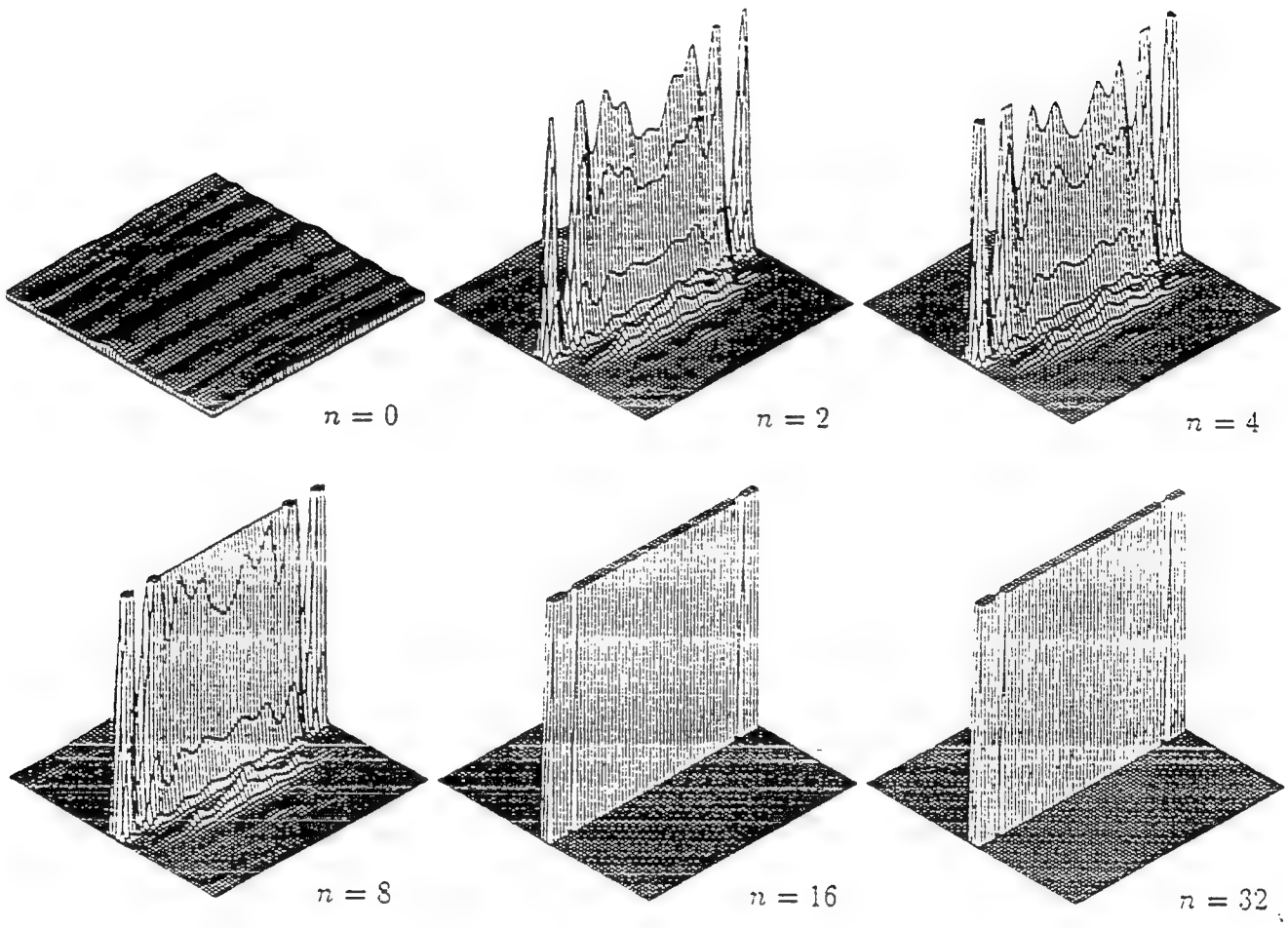



Figure 4

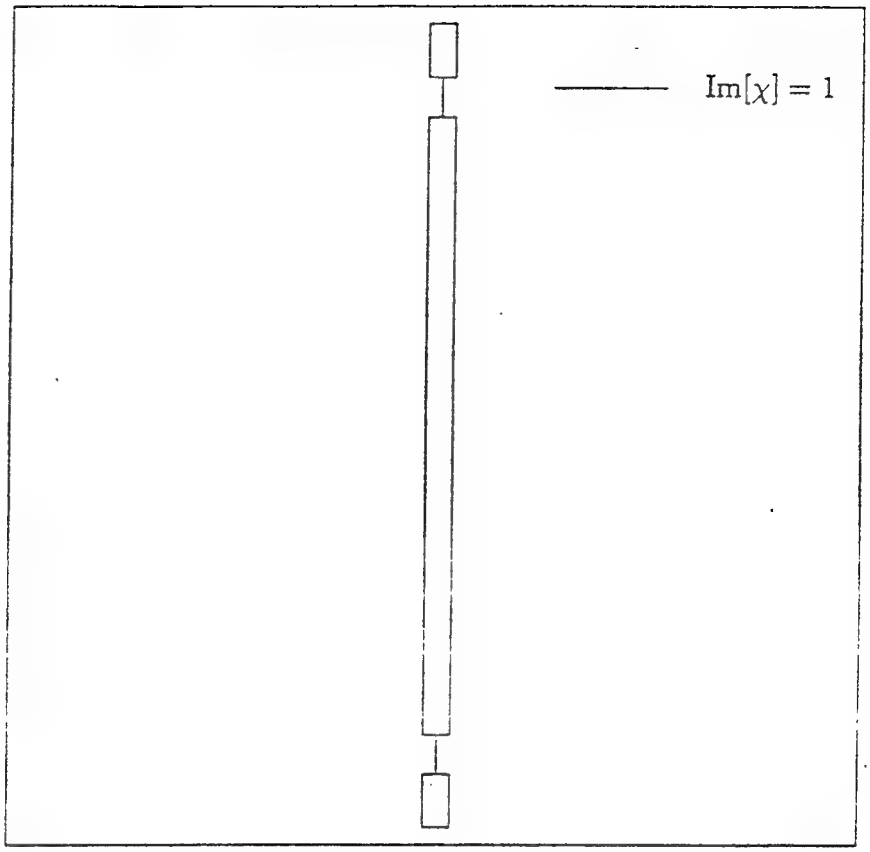



Figure 4a

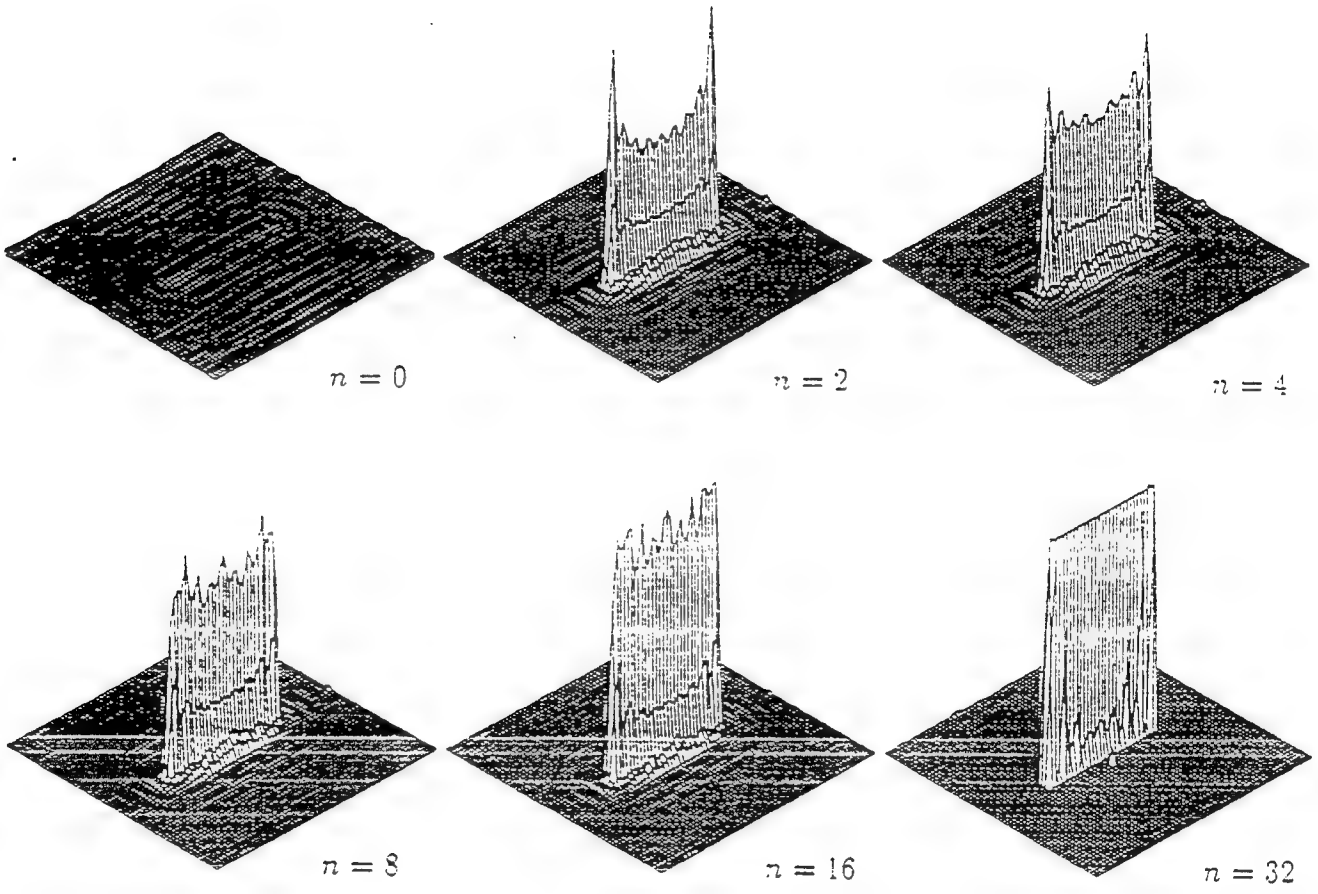



Figure 5

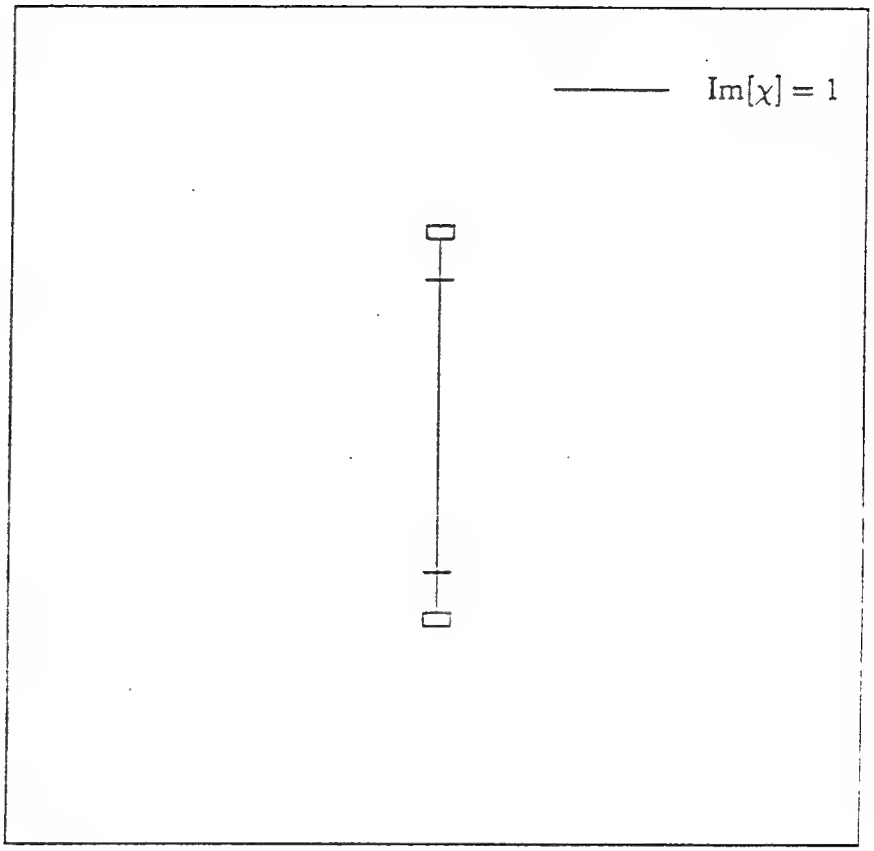
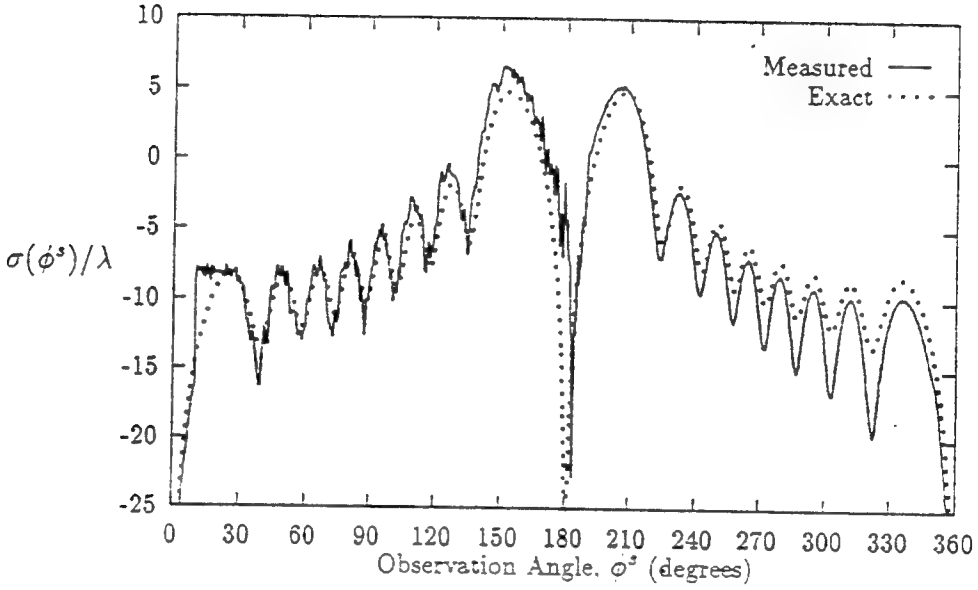



Figure 5a


Measured
Exact
$\sigma(\phi^s)/\lambda$
10
5
0
-5
-10
-15
-20
-25
0 30 60 90 120 150 180 210 240 270 300 330 360
Observation Angle, $\phi^s$ (degrees)

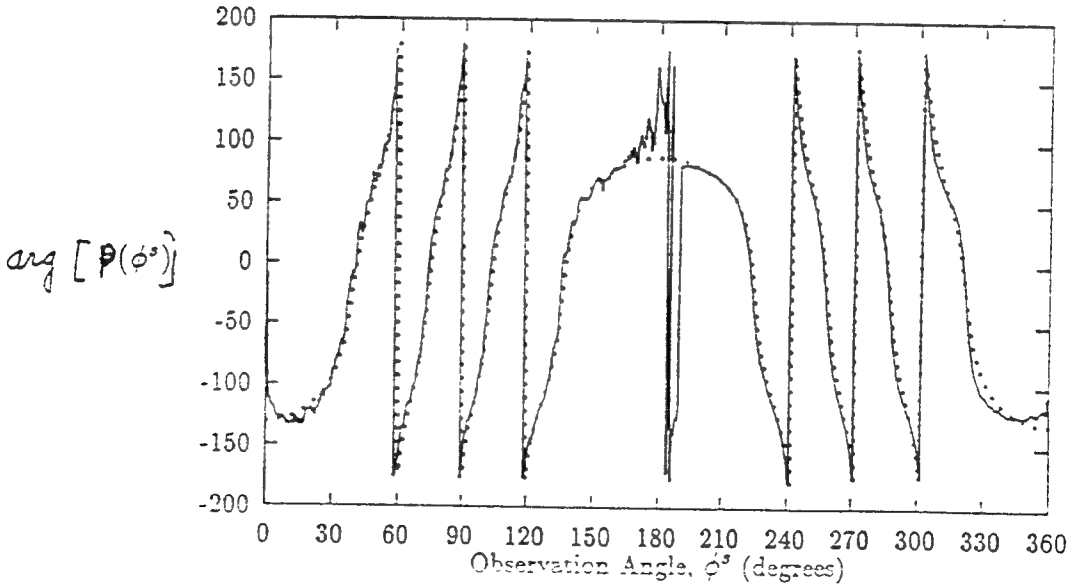

arg $[\Phi(\phi^s)]$
200
150
100
50
0
-50
-100
-150
-200
0 30 60 90 120 150 180 210 240 270 300 330 360
Observation Angle, $\phi^s$ (degrees)

**Figure 6**

# FULL LOW-FREQUENCY ASYMPTOTIC EXPANSION FOR SECOND ORDER ELLIPTIC EQUATIONS IN TWO DIMENSIONS

**Ralph Kleinman**
Department of Mathematical Sciences
University of Delaware
Newark, DE 19716

and

**Boris Vainberg**
Department of Mathematical Sciences
University of North Carolina
Charlotte, NC 28223

**Technical Report No. 93–7**

To appear in: Math. Methods in the Applied Sciences.

## Abstract

The present paper contains the low frequency expansions of solutions of a large class of exterior boundary value problems involving second order elliptic equations in two dimensions. The differential equations must coincide with the Helmholtz equation in a neighborhood of infinity, however they may depart radically from the Helmholtz equation in any bounded region provided they retain ellipticity. In some cases the asymptotic expansion has the form of a power series with respect to $k^2$ and $k^2(\ln k + a)^{-1}$ where $k$ is the wave number and $a$ is a constant. In other cases it has the form of a power series with respect to $k^2$, coefficients of which depend polynomially on $\ln k$. The procedure for determining the full low frequency expansion of solutions of the exterior Dirichlet and Neumann problems for the Helmholtz equation is included as a special case of the results presented here.

## Introduction and Formulation of the Main Results

Let $\Omega$ be an unbounded domain in $\mathbb{R}^2$ with compact infinitely smooth boundary $\Gamma$, let

$$A = \sum_{i,j=1}^{2} \frac{\partial}{\partial x_i} a_{ij}(x) \frac{\partial}{\partial x_j} + \sum_{i=1}^{2} b_i(x) \frac{\partial}{\partial x_i} + c(x)$$

be an elliptic operator of the second order (that is the matrix $(a_{ij}(x))$ is non-singular) with infinitely smooth coefficients in $\bar{\Omega}$ and $a_{ij}(x)$ real valued and let $A$ coincide with the Laplace operator $\Delta$ in some neighborhood of infinity. Denote by $u$, a solution of the problem

$$\begin{cases} Au + k^2 u = f, & x \in \Omega \\ Bu = 0, & x \in \Gamma \end{cases} \tag{1}$$

where $B$ is either the identity (Dirichlet boundary condition) or the following operator

$$Bu = \frac{\partial u}{\partial \nu} + p(x) \frac{\partial u}{\partial l} + q(x) u. \tag{2}$$

Here $\frac{\partial}{\partial \nu} = \sum_{i,j=1}^{2} a_{ij}(x) n_i \frac{\partial}{\partial x_j}$ is the derivative along the conormal vector ($n = (n_1, n_2)$ is the unit vector which is normal to $\Gamma$ and directed into $\Omega$), $\frac{\partial}{\partial l}$ is the derivative along $\Gamma$, $p, q \in C^\infty$. Finally denote by

$$R_k u : L_2(\Omega) \to H^2(\Omega), \quad \mathrm{Im} k > 0$$

the operator which takes functions $f \in L_2(\Omega)$ into solutions of problem (1) belonging to the Sobolev space $H^2(\Omega)$.

With $a$ an arbitrary constant we define a cutoff function $\chi = \chi(x) \in C^\infty(\bar{\Omega})$, such that $\chi = 1$ when $|x| < a$ and $\chi = 0$ when $|x| > a + 1$. Then a restricted

resolvent is defined as

$$\hat{R}_k := \chi R_k \chi : L_2(\Omega) \to H^2(\Omega).$$

The operators $\hat{R}_k, R_k$ are defined and are meromorphic functions of $k$ when $\mathrm{Im}k > 0$. Moreover the operator $\hat{R}_k, \mathrm{Im}k > 0$, has a meromorphic continuation on the Riemann surface of the function $\ln k$ (see [15]). Let us stress that $\hat{R}_k f = \chi R_k f$ if $f = 0$ for $|x| > a$. In this case the function $u = \hat{R}_k f$ is a solution of (1) for $|x| < a$.

The present work is devoted to the study of the asymptotic behavior of the operator $\hat{R}_k$ (that is of the solution $u = \hat{R}_k f$, $|x| < a$ of the problem (1) with $f = 0$ for $|x| > a$) as $k \to 0$. We consider only the two-dimensional case, since in other dimensions the asymptotic behavior of the solution of this problem is much simpler. For example, if the dimension is odd the solution is meromorphic in $k$ in the entire $k$ plane [see 15]. The two-dimensional problem was studied in [1] - [18]. In particular the latest results were obtained in [16]. There the problem was supposed to be formally self-adjoint (i.e. $(Au, v) = (u, Av)$ for functions $u, v \in C_0^\infty(\bar{\Omega})$ satisfying specified boundary conditions) and nonpositive, or to be more exact it was supposed that

$$(Au, u) \le -\alpha \int_\Omega |\nabla u|^2 dx, \ \alpha > 0 \quad \text{for all } u \in C_0^\infty(\bar{\Omega}) \text{ with } Bu|_\Gamma = 0. \qquad (3)$$

Here $(\cdot, \cdot)$ denotes the inner product in $L_2(\Omega)$. For this case [16] gives the asymptotic behavior of the solution $u = u_k$ of the problem (1) with accuracy $O(k^2)$.

However more than ten years ago, in [15], there appeared results of one of the present authors concerning the low-frequency asymptotic behavior of solutions

of general elliptic problems of any order polynomially depending on the spectral parameter. Those results apply to problem (1) and allow one to obtain the full asymptotic expansion of the operator $\hat{R}_k$ as $|k| - 0$. This expansion has the form

$$\hat{R}_k = k^{-2\alpha} \sum_{m=0}^{\infty} \sum_{n=0}^{m\ell} [\frac{k^2}{P(\ln k)}]^m \ln^n k \, P_{m,n}, \quad |k| << 1, \tag{4}$$

where $\alpha$ is an integer, $\ell$ is a non-negative integer, $P$ is a polynomial with constant coefficients, and $P_{m,n} : L_2(\Omega) \to H^2(\Omega)$ are bounded operators independent of $k$.

This expansion is meaningful in the sense of the operator norm, that is

$$\|\hat{R}_k f - k^{-2\alpha} \sum_{m=0}^{N} \sum_{n=0}^{m\ell} [\frac{k^2}{P(\ln k)}]^m \ln^n k P_{m,n} f\|_{H^2(\Omega)} \leq$$
$$\leq c|k^{-2\alpha+2N+2} \ln^s k| \|f\|_{L^2(\Omega)}$$

where $N$ is arbitrary, $s = \ell(N+1)$ and $c$ is independent of $k$.

The integers $\alpha$ and $\ell$ and the polynomial $P$ are not known in general, even for equations of 2nd order. It is the purpose of the present work to specify the precise form of expansion (4) for solutions of two restricted cases of problem (1).

Case I. The space of bounded solutions of the homogeneous problem

$$Au = 0, \; x \in \Omega; \; Bu = 0, \; x \in \Gamma \tag{5}$$

consists of only the trivial solution.

**Case II.** The space of bounded solutions of (5) is one dimensional and if $u$ is a nontrivial solution then

$$\lim_{r \to \infty} u(x) \neq 0. \tag{6}$$

and the formal adjoint to problem (5) (see (16) below) also has a bounded solution with property (6).

The exterior Dirichlet problem for the Helmholtz equation is an example of Case I while the exterior Neumann problem for the Helmholtz equation is an example of Case II.

It is well known that if $u$ is a bounded solution of Laplace's equation in a neighborhood of $\infty$ then for $r$ sufficiently large, $u$ has the form

$$u(x) = C_0 + \sum_{n=0}^{\infty} (a_n \cos n\varphi + b_n \sin n\varphi) r^{-n} \tag{7}$$

In particular this representation is valid for bounded solutions of problem (5) and of problem (1) with $k = 0$ if $f = 0$ in a neighborhood of infinity. Therefore, condition(6) is equivalent to the requirement that $C_0 \neq 0$ for such solutions.

The conditions embodied in Cases I and II are less restrictive than those used in [16]. If condition (3) required in [16] is fulfilled then together with (7) it follows that there are only constant solutions of problem (5). This means that either Case I or Case II apply. In our work we do not require nonpositivity, condition (3), nor do we require that the problem be self adjoint. Moreover we obtain not only the first few terms as in [16], but the complete asymptotic expansion of the solutions.

Unlike [16] we consider (for simplicity) only the problems in which the boundary and coefficients of the equation are infinitely smooth.

The main results are contained in two theorems which are presented in this paper. Let $a$ be an arbitrary fixed constant such that $\Gamma$ is contained in the circle

$|x| < a-1$ and $f = 0$ when $|x| > a$. Let $\Omega_a = \Omega \cap \{x : |x| < a\}$ and $A = \Delta$ when $|x| > a-1$. Let $L_{2,a}$ be the space of functions which belong to $L_2(\Omega)$ and are equal to zero when $|x| > a$. In particular, $f \in L_{2,a}$.

In Case I we denote by $u_0, u_1, u_2$ the solutions of the problems

$$\begin{cases} Au_0 = f, & x \in \Omega \\ Bu_0 = 0,\ x \in \Gamma; & |u_0| < \infty \text{ as } r \to \infty \end{cases} \tag{8}$$

$$\begin{cases} Au_1 = 0, & x \in \Omega; \\ Bu_1 = 0,\ x \in \Gamma; & |u_1 - \ln r| < \infty \text{ as } r \to \infty \end{cases} \tag{9}$$

$$\begin{cases} Au_2 = 0, & x \in \Omega \\ Bu_2 = 0,\ x \in \Gamma; & |u_2 + \frac{1}{2}a_1 x_1 + \frac{1}{2}b_1 x_2| < \infty \text{ as } r \to \infty \end{cases} \tag{10}$$

where $a_1, b_1$ are the coefficients in the expansion (7) for $u_0$. We will show that the uniqueness of the solution of problem (8) leads to the solvability of this problem. After we have established this, we can easily infer the unique solvability of problems (9) and (10) by reducing them to problems of the form (8). This is accomplished by writing the solutions $u_1$ and $u_2$ of problems (9) and (10) in the form $u_1 = \psi \ln r + w_1$ and $u_2 = -\frac{1}{2}(a_1 x_1 + b_1 x_2)\psi + w_2$, where $\psi = \psi(x)$ is a cutoff function which is equal to one for $|x| > a$ and equal to zero in some neighborhood of $\Gamma$. Then the problem of finding $w_i$ is of the form (8). Moreover since the expansion (7) is valid for $w_i$, in particular the constant

$$\lambda_0 = \lim_{r \to \infty} (u_1 - \ln r) \tag{11}$$

is defined. Finally, we denote by $\beta$ the constant which occurs in the asymptotic expansion of $H_0^{(1)}(z)$, the Hankel function of the first kind and order zero:

$$H_0^{(1)}(z) = \frac{2i}{\pi}(\ln z - \beta) + O(z^2 \ln z),\ z \to 0.$$

With this notation established we now state the main results.

Theorem 1. *In Case I for the solution $u = \hat{R}_k f$ of problem (1) with $f \in L_{2,a}$, the following asymptotic expansion is valid when $-\frac{\pi}{2} \leq \arg k \leq \frac{3\pi}{2}$, $|k| \to 0$ :*

$$u = \sum_{m=0}^{N} \sum_{n=0}^{m} \sum_{p=0}^{m-n+1} k^{2m} \ln^n k (\ln k - \lambda_0 - \beta)^{-p} u_{m,n,p}(x) + \tilde{u}_N \tag{12}$$

*where $u_{m,n,p}(x)$ are independent of $k$ and*

$$\|\tilde{u}_N\|_{H^2(\Omega_a)} \leq C(a)|k^2 \ln k|^{N+1} \|f\|_{L_{2,a}}. \tag{13}$$

*The leading terms of the asymptotic expansion have the form*

$$u = u_0(x) + \frac{C_0}{\ln k - \lambda_0 - \beta} u_1(x) + k^2 \ln k u_2(x) + O(k^2) \tag{14}$$

*where $u_0, u_1, u_2$ are the solutions of problems (8) - (10) and $C_0 = \lim\limits_{r \to \infty} u_0(x)$.*

Remark: In fact the corresponding expansion for the operator $\hat{R}_k$ converges in the operator norm for $0 < |k| < |k_0|$ for some $|k_0| > 0$ and therefore the infinite series for $u(N = \infty)$ converges in $H^2(\Omega_a)$.

In Case II let us denote by $v_0$ the solution of problem (5) such that

$$\lim_{r \to \infty} v_0(x) = 1. \tag{15}$$

Let the problem

$$A^* u = 0, \; x \in \Omega; \quad B^* u = 0, \quad x \in \Gamma \tag{16}$$

be formally adjoint to (5), that is, the operator $A^*$ can be obtained from $A$ by substituting $\bar{b}_i$ for $b_i$ and $\bar{c} - \sum \frac{\partial \bar{b}_i}{\partial x_i}$ for $c$. If $B = I$ then $B^* = I$, if $B$ has the form (2) then $B^*$ has the same form with $\bar{p}$ instead of $p$ and $-\frac{\partial \bar{v}}{\partial \ell} - \bar{q} + \sum \bar{b}_i(x) n_i$ instead of $q$. If $u, v \in C^\infty(\bar{\Omega})$, and $Bu = B^* v = 0$ on $\Gamma$ then

$$\int_{\Omega_R} Au\bar{v}dx = \int_{\Omega_R} u\overline{A^* v}dx + \int_{r=R} (\frac{\partial u}{\partial r}\bar{v} - u\frac{\partial \bar{v}}{\partial r})dS, \; R > a. \tag{17}$$

We will show that the space of bounded solutions of problem (16) in Case II is also one-dimensional and there exists a unique solution $v_*$ of problem (16) such that

$$\lim_{r \to \infty} v_*(x) = 1. \tag{18}$$

Let us denote by $v_1$ the solution of the inhomogeneous problem (5) (that is the solution of problem (1) with $k = 0$) such that

$$v_1 - \alpha(\ln r - \beta) \to 0 \text{ as } r \to \infty \tag{19}$$

where $\alpha = \alpha(f)$ is constant. We will show that in Case II such a solution exists, is unique and

$$\alpha = \frac{1}{2\pi} \int_{\Omega} f\bar{v}_* dx \tag{20}$$

Theorem 2. *In Case II for the solution $u = \hat{R}_k f$ of problem (1) with $f \in L_{2,a}$, the following asymptotic expansion is valid when $-\frac{\pi}{2} \leq \arg k \leq \frac{3\pi}{2}$, $|k| \to 0$ :*

$$u = \sum_{m=0}^{N} \sum_{n=0}^{2m+1} k^{2m} \ln^n k \; u_{m,n}(x) + \tilde{u}_N \tag{21}$$

*where $u_{m,n}$ are independent of $k$ and*

$$\|\tilde{u}_N\|_{H^2(\Omega_a)} \leq C(a)|k^{2N+2}\ln^{2N+3}k|\|f\|_{L_{2,a}} \tag{22}$$

*The leading terms of the asymptotic expansion have the form*

$$u = \alpha \ln k v_0(x) + v_1(x) + O(k^2\ln^3 k), \quad |k| \to 0 \tag{23}$$

*where $\alpha$ is defined in (20).*

If $v_0 \equiv 1$ then $u_{m,n} \equiv 0$ for $n > m+1$.

**Remarks.**

(1) The remark following Theorem I also applies here.

(2) It is not difficult to write out the sequence of problems similar to (8) - (10), from which we can find all the coefficients in the expansions (12) and (21).

(3) In the present work we have assumed that the data $f$ is independent of $k$. In many applications, of course, $f$ will be a known function of $k$. In such cases when $f$ may developed in a series in $k$ the present analysis will still apply. The result will be the product of the expansion of the inverse operator, $\hat{R}_k$, with the expansion of $f$.

## II. Proof of the Theorems

Let us denote by $\eta$ a particular function which is infinitely smooth, equal to zero for $|x| < a-1$ and equal to one for $|x| > a-\frac{1}{2}$. We will assume that a product of any function in $\Omega$ by $\eta$ or by a derivative of $\eta$ is defined in $\mathbb{R}^2$ and is equal to zero in $\mathbb{R}^2 \backslash \Omega$ (where $\eta = 0$). For any smooth $u$ let us denote by $g$ the following function

$$g(u) = g(u)(x) = u\Delta\eta + 2(\nabla u, \nabla\eta), \quad x \in \mathbb{R}^2. \tag{24}$$

It is obvious that the support of this function belongs to the annulus $a-1 \le |x| \le a-\frac{1}{2}$. We will denote by $*$ the convolution in $\mathbb{R}^2$.

We need the following three lemmas in order to prove Theorem 1.

**Lemma 1.** *For the solution $u = \hat{R}_k f$ of problem (1) and $x \in \mathbb{R}^2$ there are the following representations:*

$$\eta u = -\frac{i}{4} H_0^{(1)}(kr) * g(u) \text{ if } f \in L_{2,a-1}. \tag{25}$$

$$\eta u = -\frac{i}{4} H_0^{(1)}(kr) * [g(u) + \eta f] \text{ if } f \in L_{2,a}. \tag{25'}$$

**Proof.** Since $A = \Delta$ for $|x| > a-1$ we have from (1) that $(\Delta + k^2)u = f$ for $|x| > a-1$. Therefore

$$(\Delta + k^2)\eta u = g(u) + \eta f, \; x \in \mathbb{R}^2$$

where the right side has compact support and $\eta f = 0$ if $f \in L_{2,a-1}$. Representations (25), (25') follow directly. This establishes the lemma.

If we replace $H_0^{(1)}(kr)$ in (25) by its asymptotic expansion as $|k| \to 0$. for $r$ bounded we obtain the following:

Corollary. *If* $f \in L_{2,a-1}$ *and* $|k| \to 0$ *then*

$$\eta u = [\frac{1}{2\pi}(\ln k + \ln r - \beta) + O(k^2 \ln k)] * g(u) \tag{26}$$

*and also*

$$\eta u = [\frac{1}{2\pi}(\ln k + \ln r - \beta) - \frac{1}{8\pi} r^2 k^2 \ln k + O(k^2)] * g(u). \tag{27}$$

Next we establish the solvability of problem (8) which as discussed previously also ensures the solvability of problems (9) and (10). Moreover we also provide the leading term in the expansion, (4), of the solution of problem (1).

Lemma 2. *In case* I *problem* (8) *is solvable and the following expansion is valid for the solution* $u = \hat{R}_k f$ *of problem* (1) *with* $f \in L_{2,c}$

$$u = u_0 + \frac{C_0}{\ln k - \lambda_0 - \beta} u_1(x) + O(k^2 \ln^\gamma k), \; x \in \Omega_a, \; k \to 0 \tag{28}$$

*where* $u_0, u_1, \lambda_0, C_0, \beta$ *are the same as in Theorem 1 and* $\gamma$ *is an integer.*

Remark: The form (28) is based on the general expansion of the inverse operator (4) and hence the order symbol is used in the sense that

$$\|u - u_0 - \frac{C_0 u_1}{\ln k - \lambda_0 - \beta}\|_{H^2(\Omega_a)} \leq c|k^2 \ln^\gamma k| \|f\|_{L^2(\Omega_a)}.$$

The order symbol for solutions is used in the same sense throughout.

**Proof.** From (4) it follows that

$$u = k^{2s}\frac{Q(\ln k)}{L(\ln k)} + O(k^{2s+2}\ln^{\gamma} k),\ x \in \Omega_a,\ k \to 0 \tag{29}$$

where $s$ is an integer (not necessarily positive), $L$ is a polynomial with constant coefficients and $Q$ is a polynomial, not identically zero, whose coefficients are functions of $x$.

It is convenient here to assume that $f \in L_{2,a-1}$ but not $L_{2,a}$. Since $a$ can be chosen as large as we wish, this is not an additional restriction. Let us prove that $s = 0$. Taking the limit in equation (1) as $k \to 0$ we can see that $s$ cannot be positive. Let us suppose that $s < 0$. Then by equating coefficients of $k^{2s}$ in both sides of equation (1) it follows that the function $w := \frac{Q}{L}$ satisfies

$$Aw = 0,\ x \in \Omega_a;\ Bw = 0,\ x \in \Gamma. \tag{30}$$

On the other hand putting (29) in both sides of the equality (26) and equating the leading terms which contain the multiplier $k^{2s}$ we obtain

$$\eta w = \frac{1}{2\pi}(\ln k + \ln r - \beta) * g(w),\ x \in \Omega_a \tag{31}$$

Since the right side of (31) depends on values of $w$ only when $x \in \Omega_a$ formula (31) allows us to continue the function $w$ to the whole domain $\Omega$. As a result we obtain the function $w$ in $\Omega$ which satisfies (30) and

$$\eta w = \frac{1}{2\pi}(\ln k + \ln r - \beta) * g(w),\ x \in \Omega. \tag{32}$$

Since $\eta = 1$ and $g = 0$ for $r > a - \frac{1}{2}$ we have from (32) that $\Delta w = 0$ for $r > a - \frac{1}{2}$. From this and (30) it follows that $w$ is a solution of problem (5). On the other hand from (32) it follows that

$$w - c(k)(\ln k + \ln r - \beta) \to 0, \text{ as } r \to \infty \tag{33}$$

where

$$c(k) = \frac{1}{2\pi} \int_{\Omega} g(w) dx. \tag{34}$$

Since $w = \frac{Q}{L}$ for $x \in \Omega_a$ and $w$ is defined by (32) for $r > a$ there exists an integer $\nu$ and functions $w_0, w_1$ of $x$ such that $w_0 \neq 0$ identically and

$$w = w_0 \ln^{\nu} k + w_1 \ln^{\nu-1} k + O(\ln^{\nu-2} k), \; k \to 0 \tag{35}$$

Obviously $w_0$ is a solution of problem (5) because $w$ is a solution. On the other hand putting (35) in both sides of (32) and equating the leading terms (of order $\ln^{\nu+1} k$ and $\ln^{\nu} k$) of the asymptotic expansions, we obtain

$$0 = \int_{\Omega} g(w_0) dx \tag{36}$$

$$\eta w_0 = \frac{1}{2\pi} (\ln r - \beta) * g(w_0) + \frac{1}{2\pi} \int_{\Omega} g(w_1) dx \tag{37}$$

From here it follows that $w_0 \to$ constant as $r \to \infty$. Therefore $w_0$ is a bounded solution of the problem (5) and it must be zero. This contradiction proves that $s = 0$.

Having established that $s = 0$, the same argument can be repeated, with $s = 0$, to show that $w = \frac{Q}{L}, x \in \Omega_a$ can be continued to all of $\Omega$ and (32) - (37) are fulfilled.

However $w$ is now a solution of an inhomogeneous problem (problem (1) with $k = 0$). Since $w$ has the form (35) and is a solution of problem (1) with $k = 0$ we may take the limit as $k \to 0$. This process produces a contradiction if $\nu < 0$ (i.e. $0 = f$). Hence $\nu \geq 0$. If $\nu > 0$ we multipy both sides of the equations by $\ln^{-\nu} k$ and again take the limit as $k \to 0$. We find that $w_0$ is a solution of the homogeneous problem (1) with $k = 0$. In addition we may conclude that $w_0$ is bounded using the same arguments used previously (c.f. (36) and (37)). Therefore $w_0$ must be equal to zero. This shows that $\nu = 0$ and $w_0$ is a bounded solution of problem (1) with $k = 0$. This means that $w_0 = u_0$ and therefore the existence of the solution $u_0$ of problem (8) is proved, hence, as noted previously, the solution $u_1$ of problem (9) also exists.

Now recall that we have the expansion (29) with $s = 0$ in which the function $w = \frac{Q}{L}$ is a solution of problem (1) with $k = 0$ and satisfies (33) and (34). Then the function $w - c(k)u_1$ is a solution of (1) with $k = 0$ which tends to $c(k)(\ln k - \lambda_0 - \beta)$ as $r \to \infty$. It means that $w - c(k)u_1 = u_0$ and $c(k)(\ln k - \lambda_0 - \beta) = C_0$ which can be solved for $c(k)$.

Thus we have shown that

$$u = w + O(k^2 \ln^\gamma k) = u_0 + c(k)u_1 + O(k^2 \ln^\gamma k).$$

With the formula for $c(k)$, (28) is established and Lemma 2 is therefore proved.

**Remark.** Since $w_0 = u_0$ it follows from (36) that

$$\int_\Omega g(u_0) dx = 0 \qquad (38).$$

Lemma 2 gives the two leading terms of the asymptotic form of the solution of problem (1) in Case I. It is possible to obtain higher order terms in the same

way. In particular one can show that $\gamma$ in (28) is equal to one. However this result is a direct consequence of the complete expansion (12) in Theorem I which will be proven below. In order to specify the first three terms of (12) the general theorem yields a result of the form (14). To complete the proof that the coefficients in (14) are solutions of (8) - (10), we will need the following.

**Lemma 3.** *In Case* I *if the constant* $\gamma$ *in* (28) *is equal to one then the asymptotic expansion* (14) *is valid.*

**Proof.** If $\gamma = 1$ then from (4) and Lemma 2 it follows that

$$u = u_0 + \frac{C_0}{\ln k - \lambda_0 - \beta} u_1(x) + k^2 \ln k u_2(x) + k^2 u_3(x) + O(k^2 \ln^{-1} k), \quad x \in \Omega_a, \ |k| \to 0 \tag{39}$$

where $u_0$ and $u_1$ are the solutions of problems (8) and (9) and $u_2(x)$ and $u_3(x)$ are functions independent of $k$. It remains to show that $u_2$ is the solution of problem (10).

Let us substitute (39) into (1) and equate terms of order $k^2 \ln k$ to zero since there are no such terms on the right hand side. We then have

$$\begin{cases} Au_2 = 0, & x \in \Omega_a; \\ Bu_2 = 0, & x \in \Gamma. \end{cases}$$

Further as in the proof of Lemma 2, we can assume that $f \in L_{2,a-1}$. Now substitute (39) in (27) and equate the terms of equal order. In particular there is only one term of order $k^2 \ln^2 k$ and its coefficient must therefore vanish. Hence

$$0 = \int g(u_2) dx \tag{40}$$

Equating the terms with $k^2 \ln k$ we obtain

$$\eta u_2 = \frac{1}{2\pi}(\ln r - \beta) * g(u_2) - \frac{1}{8\pi} r^2 * g(u_0) + \frac{1}{2\pi} \int_{R^2} g(u_3) dx, \qquad (41)$$

As was done in the proof of Lemma 2, we can use (41) in order to continue the function $u_2$ on the entire domain $\Omega$. The extended function $u_2$ will be a solution of problem (5) and, will satisfy (41) for all $x \in \Omega$. This is proved exactly as was done in the proof of Lemma 2. To complete the proof of Lemma 3 it remains to show that $u_2$ has the same asymptotic behavior when $r \to \infty$ as in (10).

Since it follows from (40) that the first term of the right side in (41) decays as $r \to \infty$ it is enough to check that

$$|\frac{1}{4\pi} r^2 * g(u_0) - a_1 x_1 - b_1 x_2| < C, \; r \to \infty.$$

We have,

$$r^2 * g(u_0) = \int_{R^2} |x - y|^2 g(u_0) dy = r^2 \int_{R^2} g(u_0) dy +$$
$$-2x_1 \int_{R^2} y_1 g(u_0) dy - 2x_2 \int_{R^2} y_2 g(u_0) dy + \int_{R^2} |y|^2 g(u_0) dy. \qquad (42)$$

According to (38) the first term in the right side of (42) is zero. Since the last term is constant it remains to show that

$$-\frac{1}{2\pi} \int_{R^2} x_1 g(u_0) dx = a_1, \quad -\frac{1}{2\pi} \int_{R^2} x_2 g(u_0) dx = b_1. \qquad (43)$$

Since $f \in L_{2,a-1}$ and $A = \Delta$ for $r > a - 1$ it follows from (8) that $\Delta u_0 = 0$ for $r > a - 1$. Therefore (see (24))

$$g(u_0) = \Delta(\eta u_0)$$

Further, according to (7), we have

$$u_0 = C_0 + a_1 \frac{\cos\varphi}{r} + b_1 \frac{\sin\varphi}{r} + O(r^{-2}),\ r \to \infty \tag{44}$$

where $\tan\varphi = \frac{y}{x}$. Hence for $R > a$ we have, with Green's theorem,

$$\int\limits_{r\le R} x_1 g(u_0)dx = \int\limits_{r\le R} x_1 \Delta(\eta u_0)dx = \int\limits_0^{2\pi} (R^2 \cos\phi \frac{\partial u_0}{\partial r}|_{r=R} - R\cos\phi u_0|_{r=R})d\phi$$

Taking the limit as $R \to \infty$ and also taking (44) into account we obtain the first equality in (43). The second one can be established in the same way. Thus, Lemma 3 is proved.

Now we may proceed to the proof of Theorem I.

Proof of Theorem 1. Let $\zeta \in C^\infty(\Omega)$, $\zeta = 1$ for $r > a - 1$, $\zeta = 0$ in a neighborhood of $\Gamma$. Let us denote by $U$ the operator which maps the righthand side into the first two terms of the solution of problem (1),

$$U : L_{2,a}(\Omega) \to H^2_{loc}(\Omega), \quad Uf = u_0 + \frac{C_0}{\ln k - \lambda_0 - \beta} u_1, \tag{45}$$

where $u_0$ and $u_1$ are the solutions of problems (8) and (9). Further let us denote by $\Phi_k$ the following operator (parametrix of problem (1)):

$$\Phi_k : L_{2,a} \to H^2_{loc}(\Omega) \tag{46}$$

$$\Phi_k h = (Uh)(1-\eta) - \frac{i\zeta}{4}[H_0^{(1)}(kr) * (g(Uh) + \eta h)],\ h \in L_{2,a} \tag{47}$$

Let us prove that for any $h \in L_{2,a}$ function $\Phi_k h$ satisfies the following relations:

$$\begin{cases} (A + k^2)\Phi_k h = h + T_k h, & x \in \Omega \\ B\Phi_k h = 0, & x \in \Gamma \end{cases} \tag{48}$$

where $T_k h \in L_{2,a}$ and

$$\|T_k h\|_{L_{2,a}} \leq C|k^2 \ln^{\gamma+1} k| \|h\|_{L_{2,a}} \tag{49}$$

Multiplying both sides of (25′) by $\zeta$ and taking into account that $\zeta\eta = \eta$ we have

$$\eta \hat{R}_k h = -\frac{i\zeta}{4}[H_0^{(1)}(k\tau) * (g(\hat{R}_k h) + \eta h)]$$

Hence the following identity is valid

$$\hat{R}_k h \equiv (\hat{R}_k h)(1-\eta) - \frac{i\zeta}{4}[H_0^{(1)}(k\tau) * (g(\hat{R}_k h) + \eta h)] \tag{50}$$

This allows us to estimate the difference $\Phi_k h - \hat{R}_k h$. Using Lemma 2 and (24) it follows that

$$\|(\hat{R}_k h - Uh)\|_{H^2(\Omega_a)} \leq C|k^2 \ln^{\gamma} k| \|h\|_{L_{2,a}}, \quad |k| \to 0 \tag{51}$$

and

$$\|g(\hat{R}_k h) - g(Uh)\|_{L_{2,a}} \leq C|k^2 \ln^{\gamma} k| \|h\|_{L_{2,a}}, \quad |k| \to 0 \tag{52}$$

Moreover it is obvious, that for any $f \in L_{2,a}$

$$\| -\frac{i}{4} H(k\tau) * f\|_{H^2(\Omega_a)} < C|\ln k| \|f\|_{L_{2,a}}, \quad |k| \to 0 \tag{53}$$

Subtracting (50) from (47) and estimating the right side of the result with the help of (51) - (53) we obtain

$$\|\Phi_k h - \hat{R}_k h\|_{H^2(\Omega_a)} \leq C|k^2 \ln^{\gamma+1} k| \|h\|_{L_{2,a}}, \quad |k| \to 0$$

and therefore

$$\|(A+k^2)\Phi_k h - h\|_{L_2(\Omega_a)} \le C|k^2 \ln^{\gamma+1} k| \|h\|_{L_{2,a}},\ |k| \to 0 \qquad (54)$$

Further for $r > a - \frac{1}{2}$ the following relations hold

$$\eta = \zeta = 1,\ A = \Delta,\ g = 0.$$

From these relations and (47) it follows that $(A+k^2)\Phi_k h = h$ for $r > a - \frac{1}{2}$. Together with (54) it follows that $T_k h \in L_{2,a}$ and the estimate (49) is valid. Finally, since $\eta = \zeta = 0$ in a neighborhood of $\Gamma$ it follows from the boundary conditions of problems (8) and (9) that the function (47) satisfies the boundary condition (48). Thus the relations (48) and (49) are established.

Now we look for the solution $u = R_k f$ of problem (1) with $\mathrm{Im} k > 0$ in the form $u = \Phi_k h$ with an unknown function $h \in L_{2,a}$. It is obvious from its construction, (47), that $\Phi_k h \in H^2(\Omega)$ if $\mathrm{Im} k > 0$. Hence, $R_k f = \Phi_k h$ if the function $\Phi_k h$ satisfies the differential equation and the boundary condition of problem (1). According to (48) both will be satisfied if $h + T_k h = f$. Therefore

$$R_k f = \Phi_k (I + T_k)^{-1} f, \quad \mathrm{Im} k > 0$$

and hence

$$\hat{R}_k f = \chi \Phi_k (I + T_k)^{-1} f, \quad f \in L_{2,a}. \qquad (55)$$

where $\chi$ is the cutoff function introduced in the definition of $\hat{R}_k$. Let us recall that according to its definition, the operator $U$ has the following form

$$U = U_0 + \frac{C_0}{\ln k - \lambda_0 - \beta} U_1 \qquad (56)$$

where the operators $U_0 : f \to u_0$. $U_1 : f \to u_1$ transform the function $f$ into the solutions $u_0$ and $u_1$ of problems (8) and (9) and the operators $U_0$ and $U_1$ do not depend on $k$. Further from the asymptotic expansion of the Hankel function as $|k| \to 0$. which in fact is convergent, it follows that for $f \in L_{2,a}$ and $|k|$ sufficiently small

$$\zeta[H_0^{(1)}(kr) * f] = [\sum_{j=0}^{\infty} k^{2j} A_j + \ln k \sum_{j=0}^{\infty} k^{2j} B_j] f \tag{57}$$

where $A_j, B_j : L_{2,a} \to H^2_{loc}(\Omega_a)$ are bounded operators. From (56), (57) and (24) it follows that the operator $\Phi_k$, (46) and (47), has the following convergent asymptotic expansion for $|k|$ sufficiently small

$$\Phi_k = \sum_{j=0}^{\infty} k^{2j} \Phi_j^{(1)} + \ln k \sum_{j=1}^{\infty} k^{2j} \Phi_j^{(2)} + \frac{1}{\ln k - \lambda_0 - \beta} \sum_{j=0}^{\infty} k^{2j} \Phi_j^{(3)} \tag{58}$$

where $\Phi_j^s : L_{2,a} \to H^2_{loc}(\Omega)$ are bounded operators, $s = 1, 2, 3$. In (58) there is no term with coefficient $\Phi_0^{(2)}$ because

$$\Phi_0^{(2)} f = -\frac{\zeta}{2\pi} \int_{\Omega} g(u_0) dx$$

and this function is equal to zero due to (38).

Since

$$T_k = (A + k^2)\Phi_k - I : L_{2,a} \to L_{2,a}$$

with (58) and (49) it follows that, for $|k|$ sufficiently small

$$T_k = \sum_{j=1}^{\infty} k^{2j} T_j^{(1)} + \ln k \sum_{j=1}^{\infty} k^{2j} T_j^{(2)} + \frac{1}{\ln k - \lambda_0 - \beta} \sum_{j=1}^{\infty} k^{2j} T_j^{(3)}, \tag{59}$$

where $T_j^s : L_{2,a} \to L_{2,a}$ are bounded operators and $s = 1, 2, 3$. From (59) we have

$$(I + T_k)^{-1} = \sum_{i=0}^{\infty}(-T_k)^i = \sum_{m=0}^{\infty}\sum_{n+p=0}^{m} k^{2m}\ln^n k \frac{1}{(\ln k - \lambda_0 - \beta)^p} T_{m,n,p}$$

The representation of the solution, (12), the estimate of the remainder, (13), and the remark about convergence of the series are all obvious consequences of this last formula, (55), and (58). The expansion (14) follows from (12), (13) and Lemma 3. Theorem 1 is therefore proved.

Now we pass on to the proof of Theorem 2. First we establish two additional lemmas.

**Lemma 4.** *In Case II the following three assertions are valid:*

1) *The space of bounded solutions of problem* (16) *is one-dimensional.*

2) *There exists a unique solution $v_*$ of problem* (16) *for which condition* (18) *is fulfilled.*

3) *Problem (1) with $k = 0$ and $f \in L_{2,a}$ has at most one solution $v_1$ with asymptotic behavior (19) at infinity, and if it exists then the constant $\alpha$ in (19) has the form (20).*

**Proof.** Obviously we can assume that the origin of coordinates belongs to the domain $\mathbb{R}^2 \setminus \bar{\Omega}$ and there exists $r_0 > 0$ such that the circle of radius $r_0$ belongs entirely to this domain. Let $\tilde{\Omega}, \tilde{\Gamma}$ be the images of $\Omega, \Gamma$ and $\tilde{A}, \tilde{A}^*, \tilde{B}$ and $\tilde{B}^*$ be the images of $A, A^*, B$, and $B^*$ under the mapping $r \to r_1 = \frac{r_0^2}{r}$. Let

$$r_1^{-2}\tilde{A}u = 0,\ x \in \tilde{\Omega};\ \tilde{B}u = 0,\ x \in \tilde{\Gamma}; \qquad (60)$$

$$r_1^{-2}\tilde{A}^* u = 0,\ x \in \tilde{\Omega};\ \tilde{B}^* u = 0,\ x \in \tilde{\Gamma} \tag{61}$$

be the problems obtained from (5) and (16) respectively as a result of this mapping and multiplication of the equations by $r_1^{-2}$. The presence of the factor $r_1^{-2}$ allows us to prove that problems (60) and (61) are adjoint. The proof follows. Since the Jacobian of the mapping is equal to $r_1^{-2}$ it follows from (17) that

$$\int_{\tilde{\Omega}} (\tilde{A}u)\bar{v} r_1^{-2}\, dx = \int_{\tilde{\Omega}} u\overline{(\tilde{A}^* v)} r_1^{-2}\, dx \tag{62}$$

for any $u, v \in C^\infty(\overline{\tilde{\Omega}})$ such that $u = v = 0$ in some neighborhood of the origin and $\tilde{B}u = \tilde{B}^* v = 0$ on $\tilde{\Gamma}$. Since $A = A^* = \Delta$ in a neighborhood of infinity we have that $r_1^{-2}\tilde{A} = r_1^{-2}\tilde{A}^* = \Delta$ in a neighborhood of the origin and the following two assertions are valid: 1) problems (60), (61) are elliptic problems with smooth coefficients, 2) equality (62) is valid without the assumption that $u = v = 0$ in a neighborhood of the origin, that is, problems (60) and (61) are adjoint.

Since the coefficients $a_{ij}$ of the operator $A$ are real, problem (60) is homotopically equivalent to the Dirichlet problem (if $B$ is the identity operator) or Neumann problem (if $B$ has form (2)) for the Laplace operator. Therefore the index of problem (60) is zero and the dimensions of the spaces of smooth solutions of problems (60) and (61) are the same. Since the original operator $A$ and its adjoint coincide with the Laplacian in a neighborhood of infinity, inversion establishes the one-to-one correspondence of the space of bounded solutions of the exterior problems (5) and (16) with smooth solutions of the interior problems (60) and (61) respectively. Therefore the dimension of the space of bounded solutions (5) and (16) is the same. Thus, the space of the bounded solutions of problem (16) is one-dimensional (as it

is for problem (5)). The first assertion of the Lemma is therefore proved.

Since we suppose that problem (16) has a solution with property (6) (the second assumption for case II) the second assertion of Lemma 4 follows from the first one and formula (7). Finally we prove the last assertion of Lemma 4. From formula (17) for functions $v_1$ and $v_*$, we have

$$\int\limits_{\Omega_R} f\bar{v}_* dx = \int\limits_{r=R} \left( \frac{\partial v_1}{\partial r} \bar{v}_* - v_1 \frac{\partial \bar{v}_*}{\partial r} \right) dS, \; R > a \tag{63}$$

Since $\Delta v_1 = \Delta v_* = 0$ for $r > a - 1$ the expansion (7) is valid for the functions $v_1 - \alpha(\ln r - \beta)$ and $v_*$. Note that according to (19) the constant $C_0$ vanishes in the expansion (7) of the first of these functions. These facts lead to (20) if we take the limit in (63) as $R \to \infty$. The uniqueness of the solution $v_1$ is an obvious consequence of (20), (19) and (6). This completes the proof of Lemma 4.

Lemma 5. *In case* II *problem* (1) *with* $k = 0$ *and* $f \in L_{2,a}$ *has a solution* $v_1$ *with asymptotic behavior* (19) *at infinity (which is unique according to Lemma 4). The following expansion as* $k \to 0$ *is valid for the solution* $u = \hat{R}_k f$ *of problem* (1) *with* $f \in L_{2,a}$.

$$u = (\alpha \ln k) v_0 + v_1 + O(k^2 \ln^\gamma k), \; x \in \Omega_a, \; k \to 0 \tag{64}$$

*where* $\alpha$ *is given by* (20) *and* $\gamma$ *is a constant.*

Proof. From (4) it follows that

$$u = k^{2s} \frac{Q(\ln k)}{L(\ln k)} + O(k^{2s+2} \ln^\nu k), \; x \in \Omega_a, \; k \to 0, \tag{65}$$

where $Q$ and $L$ have the same form as in formula (29) and $\nu$ is some constant. As was done in the proof of Lemma 2 we deduce first that $s \leq 0$ and if $s < 0$ then the

function $w = \frac{Q}{L}$, $x \in \Omega_a$, can be continued on the whole domain $\Omega$ in such a way that $w$ is a solution of (5) and satisfies (33). Hence the difference $w - c(k)\ln kv_0$ is a solution of (5) and satisfies (19) with $\alpha = c(k)$. From the uniqueness of $v_1$ which was proved in Lemma 4 it follows that $w - c(k)\ln kv_0 \equiv 0$ and $\alpha = c(k) = 0$. Hence $w \equiv 0$. Therefore, $s = 0$. Now we can repeat all the arguments concerning $w$ and we obtain the expansion (65) with $s = 0$ in which the function $w$ is a solution of the inhomogeneous problem (1) with $k = 0$ and (33) is fulfilled. This means that the function $w - c(k)\ln kv_0$ is a solution of problem (1) with $k = 0$ which satisfies (19) with $\alpha = c(k)$. We simultaneously obtain the existence of $v_1$ and the expansion (64). Lemma 5 is thus proved.

Finally we are in a position to prove Theorem 2.

**Proof of the Theorem 2.** In order to prove the theorem we have to repeat almost word for word the proof of Theorem 1 replacing Lemma 3 by Lemma 5 and the operator $U$ by the following operator

$$V : L_{2,a}(\Omega) \to H^2_{loc}(\Omega),\ Vf = \alpha \ln k v_0 + v_1$$

where $v_0$ satisfies (5) and (15), $v_1$ satisfies (1) with $k = 0$ and (19) and $\alpha$ is given by (20). Instead of (56) we have to use the formula $V = \alpha \ln k V_0 + V_1$, where $V_0 f = v_0$, $V_1 f = v_1$. Therefore in place of (58) we obtain, for $|k|$ sufficiently small

$$\Phi_k = \sum_{j=0}^{\infty} k^{2j} \Phi_j^{(1)} + \ln k \sum_{j=0}^{\infty} k^{2j} \Phi_j^{(2)} + \ln^2 k \sum_{j=1}^{\infty} k^{2j} \Phi_j^{(3)}. \tag{66}$$

In place of (59) we have

$$T_k = \sum_{j=1}^{\infty} k^{2j} T_j^{(1)} + \ln k \sum_{j=1}^{\infty} k^{2j} T_j^{(2)} + \ln^2 k \sum_{j=1}^{\infty} k^{2j} T_j^{(3)}, \tag{67}$$

and

$$(1+T_k)^{-1} = \sum_{m=0}^{\infty} \sum_{n=0}^{2m} k^{2m} \ln^n k T_{m,n}. \tag{68}$$

Expansions (21) and (22) follow from (55), (67) and (68). Formula (23) is a consequence of (22) and Lemma 5.

In order to specify the leading terms, (21) for the case $v_0 = 1$ we must slightly change the operator $\Phi_k$. We add the term $(k^2 \ln^2 k)Q$ to the operator $\Phi_k$, (47), where

$$Qf = \frac{\alpha(\zeta - 1)}{8\pi}[r^2 * g(v_0)], \; v_0 \equiv 1.$$

and $\alpha$ has form (20). It follows from (24) that $g(v_0) = \Delta\eta = \Delta(\eta - 1)$, where $\eta - 1 \in C_0^\infty$. Hence

$$Qf = \frac{\alpha(\zeta - 1)}{8\pi}[r^2 * \Delta(\eta - 1)] = \frac{\alpha(\zeta - 1)}{8\pi}[\Delta r^2 * (\eta - 1)] = const \cdot (\zeta - 1)$$

Hence the function $Qf$ is constant in a neighborhood of $\Gamma$. Since $Qf$ is a multiple of $v_0$ in a neighborhood of the boundary and $v_0$ satisfies the boundary condition of problem (1), $Qf$ also satisfies this condition. Therefore the addition of the term $k^2 \ln^2 k Q$ to $\Phi_k$ does not change of the steps of the proof of the theorem, but it does change the coefficients in the expansions (66) and (67). In particular it is not difficult to check that the operator $T_1^{(3)}$ which defined the leading term in (67) and had the form

$$T_1^{(3)} f = \Delta\{-\frac{\alpha\zeta}{8\pi}[r^2 * g(v_0)]\}$$

now has the form

$$T_1^{(3)} f = -\frac{\alpha}{8\pi}\Delta[r^2 * g(v_0)], \; v_0 = 1.$$

But it was shown that the convolution in the last formula is constant. Hence the operator $T_1^{(3)}$ is equal to zero and instead of (68) we have

$$(I + T_k)^{-1} = \sum_{m=0}^{\infty} \sum_{n=0}^{m} k^{2m} \ln^n k T_{m,n}.$$

This leads to the desired form of the expansion (21). Theorem 2 is therefore proved.

**Acknowledgment:** This work was partially supported under AFOSR Grant 91-0277.

## References

[1] Eidus, D. M., 'The principle of limiting amplitude', *Russ. Math. Surv.*, 24, 97-167 (1969).

[2] Hariharan, S. I. and MacCamy, R. C., 'Low frequency acoustic and electromagnetic scattering', *Applied Numerical Mathematics*, 2, 23-35 (1986).

[3] Kress, R., 'On the limiting behavior of solutions to the boundary integral equations associated with time harmonic wave equations for small frequencies', *Math. Meth. in the Appl. Sci.*, 1, 89-100 (1979).

[4] Kress, R., 'On the low wave number asymptotics for the two-dimensional exterior Dirichlet problem for the reduced wave equation', *Math. Meth. in the Appl. Sci.*, 9, 335-341 (1987).

[5] MacCamy, R. C., 'Low frequency acoustic oscillations', *Quart. Appl. Math.*, 23, 247-256 (1965).

[6] Muravei, L. A., 'Asymptotic behavior for large values of time of the solutions of the second and third exterior boundary value problems for the wave equation with two spatial variables', *Proc. Steklov Inst. Math.*, 126, 77-158 (1973).

[7] Muravei, L. A., 'On the asymptotic behavior, for large values of time, of solutions of exterior boundary value problems for the wave equation with two space variables', *Math. USSR Sb.*, 35, 377-423 (1979).

[8] Picard, R., ' The frequency limit for time-harmonic acoustic waves', *Math. Meth. in the Appl. Sci.*, 8, 436-450 (1986).

[9] Ramm, A. G., Scattering by Obstacles, Reidel, Dordrecht (1986).

[10] Ramm, A. G., 'Behavior of the solutions to exterior boundary value problems at low frequencies', *J. Math. Anal. Appl.*, 117, 561-569 (1986).

[11] Stuwe, H. C. and Werner, P., 'Remarks on the low frequency asymptotics for the reduced wave equation and the Schrodinger equation in low-dimensional spaces', *Asymptotic Analysis*, **2**, 179-202 (1989).

[12] Tsuisumi. Y., 'Local energy decay of solutions to the free Schrodinger equation in exterior domains', *J. Fac. Sci. Univ. Tokyo*, Sect. IA, Math. **31**, 97-108 (1984).

[13] Vainberg, B. R., 'On the analytic properties of the resolvent for a certain class of operator pencils', *Math. USSR Sb.* **6**, 241-272 (1968).

[14] Vainberg, B. R., 'On exterior elliptic problems polynomially depending on a spectral parameter, and the asymptotic behavior for large time of solutions of nonstationary problems', *Math. USSR Sb.*, **21**, 221-239 (1973).

[15] Vainberg, B. R., *Asymptotic Methods in Equations of Mathematical Physics*, Moscow State Univ. Publishers, (in Russian) (1982); English Translation, Gordon and Breach Publishers, 1989.

[16] Weck, H. and Witsch, K. J., *Exact Low Frequency Analysis for a Class of Exterior Boundary Value Problems for the Reduced Wave Equations in Two Dimensions*, J. Diff. Equs. 100, 312-340, (1992).

[17] Werner, P., 'Zur asymptotik der wellengleichung und der warmeleitungsgleichung in zweidimensionalen aubenraumen', *Math. Meth. in the Appl. Sci.*, **7**, 170-201 (1985).

[18] Werner, P., 'Low frequency asymptotics for the reduced wave equation in two-dimensional exterior spaces', *Math. Meth. in the Appl. Sci.*, 8, 134-156 (1986).

*Journal of Electromagnetic Waves and Applications, Vol. 8, No. 3, 295-313, 1994*


# Low-frequency image theory for the dielectric sphere

I. V. Lindell and J. C-E. Sten

Electromagnetics Laboratory
Helsinki University of Technology
Otakaari 5A, Espoo, Finland 02150

R. E. Kleinman

Department of Mathematical Sciences
University of Delaware
Newark, DE 19716, USA

**Abstract**–Kelvin's well-known and Neumann's lesser known image theories for the sphere, valid for static sources outside the sphere, are extended to low-frequency current sources involving a nondispersive dielectric, by expressing known field integrals as arising from suitable image currents. The image of a radial current element is seen to consist of a radial line current between the center of the sphere and the Kelvin point plus a dipole at the Kelvin point. The image of a transverse current element is a combination of a transverse current strip plus a radial bifilar line current between the center and the Kelvin point. These image currents can be interpreted as image charges of the corresponding static problem in harmonic motion. The theory is tested by known limiting cases.

## 1. INTRODUCTION

The well-known image theory of electrostatic charges outside a perfectly conducting sphere was originally introduced by William Thomson (later Lord Kelvin) in one of his first studies as a young scientist in 1845 [1]. Kelvin's image theory has since then been applied to problems of electrostatics, magnetostatics, and DC current problems involving perfectly electrically or magnetically conducting bodies with most recent contributions published in the present decade [2,3].

Extension of Kelvin's theory to material spheres did not seem to have had a successful solution before a paper [4] of 1992 by one of the present authors for the dielectric sphere in electrostatics. However, it was recently found that a similar solution was given already in 1883 by Carl Neumann in an appendix of a book on hydrodynamics [10]. Obviously, the solution has been dormant for more than a century and most probably was never elaborated beyond its introduction on a couple of pages. The solution will be referred to as 'the Neumann image' from now on. After its rediscovery, the theory has been extended to magnetostatics [5], layered dielectric spheres [6], and two separate spheres [7]. Also, chiral and bi-isotropic spheres have been solved [8] in terms of image theory as well as sources inside the sphere [9].

In the electrostatic image theory of dielectric sphere the image of a point charge was seen to consist of a point charge plus a line charge with a simple analytic power-law expression, readily accessible to small-scale computation [4]. In terms of this theory, many problems involving spheres can be formulated in a simple manner with line image sources taking the place of Huygens' surface sources or volume polarization sources in integral equations.

In the present paper, an attempt is made to extend this static theory to low-frequency problems involving a dielectric sphere. It is seen that, for a current dipole outside the sphere, an image current source can be found in the form of a line current and a bifilar line current. It is assumed that such a theory can be utilized for time-harmonic problems where the basic static approximation is not good enough, i.e., when the radius of the sphere cannot be considered to be very small in terms of the wavelength. Further extension to spheres with lossy dielectric and/or magnetic permeability seems well within reach with the method given in this paper.

## 2. THEORY

We consider the electromagnetic problem of an infinitesimal time-harmonic current dipole outside a dielectric sphere centered at the origin, Fig. 1. The radius of the sphere is $a$ and its permittivity $\epsilon_r\epsilon_o$, and the permittivity of the space outside the sphere is assumed to be $\epsilon_o$ without loss of generality. The permeability both inside and outside the sphere is assumed to be $\mu_o$. For simplicity, the sphere material is assumed to be dispersionless. The current dipole lies at the point $\mathbf{r}'$ with $|\mathbf{r}'| > a$ and it is represented by the current function

$$\mathbf{J}(\mathbf{r}) = \mathbf{u}IL\delta(\mathbf{r} - \mathbf{r}'), \tag{1}$$

where the unit vector $\mathbf{u}$ gives the direction of the dipole and $IL$ its moment. Another way to describe the dipole is in terms of the dipole moment vector $\mathbf{p}$,

$$\mathbf{J}(\mathbf{r}) = j\omega\mathbf{p}\delta(\mathbf{r} - \mathbf{r}'), \tag{2}$$

whose relation to the current function is $\mathbf{p} = \mathbf{u}IL/j\omega$.

The electric charge at the ends of the dipole is represented by the charge density function

$$\varrho(\mathbf{r}) = \frac{\nabla \cdot \mathbf{J}(\mathbf{r})}{-j\omega} = -\frac{IL}{j\omega}\mathbf{u} \cdot \nabla\delta(\mathbf{r} - \mathbf{r}') = -\mathbf{p} \cdot \nabla\delta(\mathbf{r} - \mathbf{r}') \tag{3}$$

At some stage we define the dipole to be on the cartesian coordinate $z$ axis to be able to compare with previous results. The objective is to find the image source which replaces the dielectric sphere in giving the reflected field in low-frequency approximation.

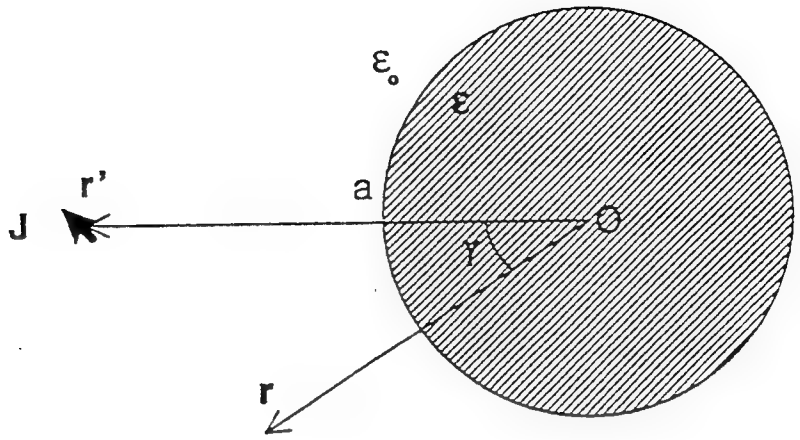



**Figure 1.** Geometry of the problem. The dielectric sphere of radius $a$ and permittivity $\epsilon = \epsilon_r \epsilon_o$ is small in terms of the free-space wavelength. The current dipole is outside the sphere at the point denoted by the vector $\mathbf{r}'$.

### 2.1 Stevenson Analysis

Let us apply the Stevenson method [12] and expand all quantities in Taylor series of $\omega$. Assuming $\epsilon$ independent of frequency, from the Maxwell equations

$$\nabla \times \mathbf{E} = -j\omega\mu_o \mathbf{H}, \tag{4}$$

$$\nabla \times \mathbf{H} = j\omega\epsilon \mathbf{E} + \mathbf{J}, \tag{5}$$

$$\nabla \cdot \mathbf{D} = \nabla \cdot (\epsilon \mathbf{E}) = \varrho, \tag{6}$$

$$\nabla \cdot \mathbf{B} = \mu_o \nabla \cdot \mathbf{H} = 0, \tag{7}$$

we can see that the equations are satisfied if $\mathbf{E}$, $\mathbf{D}$, and $\varrho$ are assumed to be even, and $\mathbf{H}$, $\mathbf{B}$, and $\mathbf{J}$ odd, functions of the frequency $\omega$. We define the order of different terms by the power of $\omega$ or, what is equivalent, of $k_o = \omega\sqrt{\mu_o \epsilon_o}$, and denote the order by the corresponding subindex. Let us assume that the current is of first order, $\mathbf{J} = \mathbf{J}_1$, i.e., a linear function of $\omega$. From (3) we see that the charge must then be of zeroth order, $\varrho = \varrho_0$.

Thus, in the static limit $\omega \to 0$, the charges are frozen and the current is zero. The resulting electrostatic equations are

$$\nabla \times \mathbf{E}_0 = 0, \tag{8}$$

$$\nabla \cdot (\epsilon \mathbf{E}_0) = \varrho_0 \tag{9}$$

plus regularity conditions at infinity. The problem of the sphere was solved in terms of electrostatic images in [4].

The first correction for the electrostatic field arises from the first-order magnetic field $\mathbf{H}_1$, whose sources are the first-order current $\mathbf{J}_1$ and the zeroth-order electric field satisfying

$$\nabla \times \mathbf{H}_1 = j\omega\epsilon\mathbf{E}_0 + \mathbf{J}_1, \tag{10}$$

$$\nabla \cdot \mathbf{H}_1 = 0. \tag{11}$$

The next set of equations reads

$$\nabla \times \mathbf{E}_2 = -j\omega\mu_o\mathbf{H}_1, \tag{12}$$

$$\nabla \cdot (\epsilon \mathbf{E}_2) = 0, \tag{13}$$

and so on. The problem considered here is to extend the zeroth-order electrostatic image theory for the first-order magnetic field by finding the first-order image currents.

### 2.2 Zeroth-Order Problem

The zeroth-order electric field can be written in terms of a scalar potential:

$$\mathbf{E}_0 = -\nabla\phi_0, \tag{14}$$

which satisfies the Poisson equation (9)

$$\nabla \cdot (\epsilon \nabla \phi_0) = -\varrho_0. \tag{15}$$

This equation can also be written as

$$\nabla^2\phi_0 = -\frac{\varrho_0}{\epsilon_o} + \frac{\nabla \cdot \mathbf{P}_0}{\epsilon_o}, \tag{16}$$

where $\mathbf{P}_0$ denotes the secondary source, the polarization moment density of the dielectric sphere:

$$\mathbf{P}_0 = (\epsilon_r - 1)\epsilon_o\mathbf{E}_0. \tag{17}$$

In (16), the primary and secondary sources can be defined to give rise to the incident and scattered potentials, $\phi_0 = \phi_0^{inc} + \phi_0^s$, in the region outside the sphere, satisfying

$$\nabla^2 \phi_0^{inc} = -\frac{\varrho_0}{\epsilon_o}, \tag{18}$$

$$\nabla^2 \phi_0^r = \frac{\nabla \cdot \mathbf{P}_0}{\epsilon_o}. \tag{19}$$

It was seen in [4] that the latter source can be replaced by a simpler image source $\varrho_{i0}$ whose expression depends on the original source. The corresponding potential satisfies

$$\nabla^2 \phi_{i0}^r = -\frac{\varrho_{i0}}{\epsilon_o}, \tag{20}$$

The image source is chosen so that the difference $\phi_0^r(\mathbf{r}) - \phi_{i0}(\mathbf{r})$ vanishes outside the sphere, i.e., $-\nabla \cdot \mathbf{P}_0 - \varrho_{i0}$ is a non-radiating source.

### 2.3 First-Order Problem

Because of (11), the first-order magnetic field can be derived from a vector potential $\mathbf{A}_1$:

$$\mathbf{H}_1 = \frac{1}{\mu_o} \nabla \times \mathbf{A}_1, \tag{21}$$

whose equation can be written from (10) as

$$\nabla \times (\nabla \times \mathbf{A}_1) = -\nabla^2 \mathbf{A}_1 + \nabla(\nabla \cdot \mathbf{A}_1) = \mu_o \mathbf{J}_1 + j\omega\mu_o \mathbf{P}_0 - \nabla(j\omega\mu_o\epsilon_o\phi_0). \tag{22}$$

We are free to choose one scalar condition for $\mathbf{A}_1$. For the Lorenz condition

$$\nabla \cdot \mathbf{A}_1 = -j\omega\mu_o\epsilon_o\phi_0 \tag{23}$$

(22) is simplified. Splitting the vector potential into incident and reflected parts $\mathbf{A}_1 = \mathbf{A}_1^{inc} + \mathbf{A}_1^r$ we can write

$$\nabla^2 \mathbf{A}_1^{inc} = -\mu_o \mathbf{J}_1, \tag{24}$$

$$\nabla^2 \mathbf{A}_1^r = -j\omega\mu_o \mathbf{P}_0. \tag{25}$$

Thus, the first-order reflected magnetic field is due to the equivalent volume current $j\omega\mathbf{P}_0$ within the sphere, and it is known if the zeroth-order problem has been solved. However, we wish to simplify the volume source by replacing it through a simpler image source $\mathbf{J}_{i1}$.

## 3. LF IMAGE EXPRESSIONS

To find the image sources $\varrho_{i0}$ and $\mathbf{J}_{i1}$, we write the low-frequency approximation of a known expression for the exact reflected electric field due to the dipole source $\mathbf{J}(\mathbf{r}) = j\omega\mathbf{p}_0\delta(\mathbf{r} - \mathbf{r}')$ and try to identify its sources.

### 3.1 Green Dyadic

The exact expression for the reflected electric field can be found, e.g., from the monograph by Jones [13], and it can be written with some change in notation as

$$\mathbf{E}^r(\mathbf{r}) = -j\omega\mu_o \overline{\overline{G}}^r(\mathbf{r},\mathbf{r}') \cdot \mathbf{J}_1(\mathbf{r}') = \frac{k_o^2}{\epsilon_o} \overline{\overline{G}}^r(\mathbf{r},\mathbf{r}') \cdot \mathbf{p}_0. \tag{26}$$

The exact reflection Green dyadic is

$$\begin{aligned}\overline{\overline{G}}^r(\mathbf{r},\mathbf{r}') =& \frac{j}{4\pi k_o}\left(k_o^2 \mathbf{u}_r + \nabla\frac{\partial}{\partial r}\right)\left(k_o^2 \mathbf{u}_r' + \nabla'\frac{\partial}{\partial r'}\right)\left[rr' V_\epsilon(\mathbf{r},\mathbf{r}')\right] \\ &+ \frac{jk_o}{4\pi}(\mathbf{r}\times\nabla)(\mathbf{r}'\times\nabla') V_\mu(\mathbf{r},\mathbf{r}'),\end{aligned} \tag{27}$$

where $\mathbf{r}$ is the field point and $\mathbf{r}'$, the source point.

The potential function $V_\tau$ for $\tau = \epsilon$ and $\tau = \mu$ can be expressed as series involving spherical Hankel functions and Legendre polynomials. Low-frequency approximation up to the first order in $k_o$ can be written, when $k_o r$ and $\sqrt{\epsilon_r} k_o r$ are small, in a simple form obtainable from the exact expressions given by Jones [13]. After some algebra,

$$V_\tau(\mathbf{r},\mathbf{r}') = \sum_{n=1}^{\infty} \frac{j}{nk_o a} \frac{\tau - 1}{(\tau+1)n+1}\left(\frac{a^2}{rr'}\right)^{n+1} P_n(\cos\gamma), \tag{28}$$

with

$$\cos\gamma = \frac{\mathbf{r}\cdot\mathbf{r}'}{rr'}. \tag{29}$$

Note that the angle $\gamma$ does not depend on the distances $r$ and $r'$. Thus, we can write

$$\frac{\partial^2}{\partial r \partial r'}(rr' V_\epsilon) = \sum_{n=0}^{\infty} \frac{j}{k_o a} \frac{n(\epsilon_r - 1)}{n(\epsilon_r+1)+1}\left(\frac{a^2}{rr'}\right)^{n+1} P_n(\cos\gamma). \tag{30}$$

### 3.2 Zeroth-Order Image

Assuming $\mu_r = 1$ for the dielectric sphere, we have $V_\mu = 0$ and the zeroth-order problem comes from the basic term and its solution should coincide with that derived earlier in [4]. From (27) and (28) we can approximate the Green dyadic by its lowest-order term, which turns out to have the order $-2$ :

$$\begin{aligned}\overline{\overline{G}}^r(\mathbf{r},\mathbf{r}') \approx \overline{\overline{G}}^r_{-2}(\mathbf{r},\mathbf{r}') &= \frac{j}{4\pi k_o}\nabla\nabla'\frac{\partial^2}{\partial r\partial r'}(rr' V_\epsilon) \\ &= -\frac{1}{k_o^2}\nabla\nabla' \sum_{n=0}^{\infty} \frac{1}{4\pi a} \frac{n(\epsilon_r-1)}{n(\epsilon_r+1)+1}\left(\frac{a^2}{rr'}\right)^{n+1} P_n(\cos\gamma).\end{aligned} \tag{31}$$

At this point, assuming $r' > a$, we write in analogy to [4]

$$\frac{n(\epsilon_r - 1)}{n(\epsilon_r + 1) + 1} \left(\frac{a^2}{r'}\right)^{n+1} = a \int_0^a f(r', r'')(r'')^n dr'', \tag{32}$$

where the image function $f(r', r'')$ is defined as

$$f(r', r'') = -\frac{d}{dr''} \left[ \frac{\epsilon_r - 1}{\epsilon_r + 1} \frac{a}{r'} \left(\frac{r'r''}{a^2}\right)^{\frac{1}{\epsilon_r + 1}} U\left(\frac{a^2}{r'} - r''\right) \right], \tag{33}$$

and $U(x)$ denotes the Heaviside unit step function. Equation (32) can be easily verified by direct substitution of (33).

Inserting (32) in (31) gives us the representation

$$\overline{\overline{G}}_{-2}^{r}(\mathbf{r}, \mathbf{r}') = -\frac{1}{4\pi k_o^2} \nabla\nabla' \int_0^a f(r', r'') \sum_{n=0}^{\infty} \frac{(r'')^n}{r^{n+1}} P_n(\cos\gamma) dr''. \tag{34}$$

Defining the vector $\mathbf{r}''$ to be in the same direction as $\mathbf{r}'$:

$$\mathbf{r}'' = \frac{\mathbf{r}'}{r'} r'' \quad \rightarrow \quad \cos\gamma = \frac{\mathbf{r} \cdot \mathbf{r}''}{r r''}, \tag{35}$$

and the distance function $D$ as

$$D(\mathbf{r} - \mathbf{r}'') = \sqrt{(\mathbf{r} - \mathbf{r}'') \cdot (\mathbf{r} - \mathbf{r}'')} = \sqrt{r^2 + (r'')^2 - 2rr''\cos\gamma}, \tag{36}$$

we can write from the definition of the Legendre polynomial for $r > r''$,

$$\frac{1}{D(\mathbf{r} - \mathbf{r}'')} = \sum_{n=0}^{\infty} \frac{(r'')^n}{r^{n+1}} P_n(\cos\gamma), \tag{37}$$

which inserted in the Green dyadic expression (34) gives us

$$\overline{\overline{G}}_{-2}^{r}(\mathbf{r}, \mathbf{r}') = -\frac{1}{k_o^2} \nabla\nabla' \int_0^a \frac{f(r', r'')}{4\pi D(\mathbf{r} - \mathbf{r}'')} dr''. \tag{38}$$

Note that it is not only $f(r', r'')$ which depends on $r'$, but also the distance function $D(\mathbf{r} - \mathbf{r}'')$, because $\mathbf{r}'' = \mathbf{u}_r' r''$, where $\mathbf{u}_r' = \mathbf{r}'/r'$ is the unit vector of both $\mathbf{r}'$ and $\mathbf{r}''$. In fact, we can write

$$\nabla' \frac{1}{D(\mathbf{r} - \mathbf{u}_r' r'')} = \frac{r''(\nabla' \mathbf{u}_r').(\mathbf{r} - \mathbf{u}_r' r'')}{D^3} = \frac{(\overline{\overline{I}} - \mathbf{u}_r' \mathbf{u}_r') \cdot \mathbf{r}\mathbf{r}''}{D^3 r'}, \tag{39}$$

which is a vector orthogonal to $\mathbf{r}'$ and $\mathbf{r}''$. Note that the order of differentiation and integration can be interchanged because the field is always computed outside

the image source, i.e., outside the sphere, whence the Green function always remains finite.

The zeroth-order reflected electric field can be written in the form

$$\mathbf{E}_0^r(\mathbf{r}) = \frac{k_o^2}{\epsilon_o} \overline{\overline{G}}_{-2}^r \cdot \mathbf{p}_0 = -\nabla \phi_0^r(\mathbf{r}), \tag{40}$$

$$\phi_0^r(\mathbf{r}) = \int_0^a \frac{\mathbf{p}_0 \cdot \nabla' f(r', r'')}{4\pi\epsilon_o D(\mathbf{r} - \mathbf{r}'')} dr'' + \int_0^a \frac{\mathbf{p}_0 \cdot (\overline{\overline{I}} - \mathbf{u}_r' \mathbf{u}_r') \cdot \mathbf{r} r'' f(r', r'')}{4\pi\epsilon_o r' D^3(\mathbf{r} - \mathbf{r}'')} dr''. \tag{41}$$

These expressions can be compared with those of the scalar potential due to a line charge $\varrho_{i0}(r'')$ extending from the center to the surface of the sphere,

$$\phi_\varrho(\mathbf{r}) = \int_0^a \frac{\varrho_{i0}(r'')}{4\pi\epsilon_o D(\mathbf{r} - \mathbf{r}'')} dr'', \tag{42}$$

and due to a dipole line charge of moment density $\mathbf{p}_{i0}(r'')$,

$$\phi_p(\mathbf{r}) = -\int_0^a \frac{\mathbf{p}_{i0}(r'') \cdot (\mathbf{r} - \mathbf{r}'')}{4\pi\epsilon_o D^3(\mathbf{r} - \mathbf{r}'')} dr''. \tag{43}$$

Comparing (41), (42), and (43), it is seen that the first integral of (41) corresponds to a line charge image whereas the second integral is of the form of a line dipole image whose distributed dipole moment vector is perpendicular to the line. Obviously, the former is generated by the radial component of the original dipole, whereas the latter is due to the transverse component of $\mathbf{p}_0$. Let us consider the two cases, $\mathbf{p}_0$ parallel and perpendicular to $\mathbf{r}'$, i.e., radial and transverse dipoles.

### Radial dipole

When $\mathbf{p}_0$ is parallel to $\mathbf{r}'$, the second integral of (41) vanishes and the first integral can be identified with (42) with the image line charge defined as

$$\begin{aligned} \varrho_{i0}(r) &= \mathbf{p}_0 \cdot \nabla' f(r', r) = p_0 \frac{d}{dr'} f(r', r) \\ &= -p_0 \frac{\partial^2}{\partial r' \partial r} \left[ \frac{\epsilon_r - 1}{\epsilon_r + 1} \frac{a}{r'} \left( \frac{r' r}{a^2} \right)^{\frac{1}{\epsilon_r + 1}} U\left( \frac{a^2}{r'} - r \right) \right]. \end{aligned} \tag{44}$$

By changing the spherical coordinate $r, r'$ to cartesian coordinate $z, z'$ this result can be compared with Eq. (23) of [4], obtained through another method. After making the two differentiations, the expressions can be seen to coincide, if the different definition of the dipole moment is taken into account.

**Transverse dipole**

For the transverse dipole satisfying $\mathbf{p}_0 \cdot \mathbf{r}' = 0$, we have $\mathbf{p}_0 \cdot \nabla' f(r', r'') = 0$, whence only the second integral of (41) survives and, identifying it with (43), the dipole line image function can be written as

$$\mathbf{p}_{i0}(r'') = -\mathbf{p}_0 \frac{r''}{r'} f(r', r'') = \mathbf{p}_0 \frac{ar''}{(r')^2} \frac{d}{dr''} \left[ \frac{\epsilon_r - 1}{\epsilon_r + 1} \left( \frac{r'r''}{a^2} \right)^{\frac{1}{\epsilon_r + 1}} U\left( \frac{a^2}{r'} - r'' \right) \right]$$

$$= -\mathbf{p}_0 \frac{\epsilon_r - 1}{\epsilon_r + 1} \left[ \left( \frac{a}{r'} \right)^3 \delta(r'' - \frac{a^2}{r'}) - \frac{1}{\epsilon_r + 1} \frac{a}{(r')^2} \left( \frac{r'r''}{a^2} \right)^{\frac{1}{\epsilon_r + 1}} U\left( \frac{a^2}{r'} - r'' \right) \right]. \tag{45}$$

Again, after a change of radial and cartesian coordinates, the result can be seen to coincide with Eq. (22) of [4].

### 3.3 First-Order Image

The reflected magnetic field $\mathbf{H}^r$ can be written in terms of the curl of the electric field as

$$\mathbf{H}^r = \frac{\nabla \times \mathbf{E}^r}{-j\omega\mu_o}. \tag{46}$$

The first-order magnetic field is written in terms of the second-order electric field:

$$\mathbf{H}_1^r = \frac{\nabla \times \mathbf{E}_2^r}{-j\omega\mu_o}. \tag{47}$$

Actually, knowledge of $\mathbf{E}_2^r$ is not needed, because $\nabla \times \mathbf{E}_2^r$ can be expressed in terms of the curl of the zeroth-order Green dyadic as

$$\mathbf{H}_1^r = -\frac{k_o^2}{j\omega\mu_o\epsilon_o} \left[ \nabla \times \overline{\overline{G}}^r \right]_0 \cdot \mathbf{p}_0 = j\omega \left[ \nabla \times \overline{\overline{G}}^r \right]_0 \cdot \mathbf{p}_0. \tag{48}$$

The zeroth-order term of the curl of the reflection Green dyadic for a nonmagnetic sphere with $\mu_r = 1$ can be written from (27), after some steps, as

$$\left[ \nabla \times \overline{\overline{G}}^r \right]_0 = -\frac{jk_o}{4\pi} (\mathbf{r} \times \nabla) \nabla' \frac{\partial}{\partial r'} (r' V_\epsilon), \tag{49}$$

$$\frac{\partial}{\partial r'} (r' V_\epsilon) = -\sum_{n=1}^{\infty} \frac{j}{k_o a} \frac{\epsilon_r - 1}{n(\epsilon_r + 1) + 1} \left( \frac{a^2}{rr'} \right)^{n+1} P_n(\cos\gamma). \tag{50}$$

Proceeding as in the previous section, we can write

$$\frac{\epsilon_r - 1}{n(\epsilon_r + 1) + 1} \left( \frac{a^2}{r'} \right)^{n+1} = a \int_0^a g(r', r'')(r'')^n dr'', \tag{51}$$

by defining

$$g(r',r'') = \frac{\epsilon_r - 1}{\epsilon_r + 1}\frac{1}{a}\left(\frac{r'r''}{a^2}\right)^{-\frac{\epsilon_r}{\epsilon_r+1}} U\left(\frac{a^2}{r'} - r''\right). \tag{52}$$

The $g(r',r'')$ function has a simple relation to the $f(r',r'')$ function defined in (33):

$$f(r',r'') = -\frac{d}{dr''}[r''g(r',r'')]. \tag{53}$$

The curl term has now the form

$$[\nabla \times \overline{\overline{G}}^r]_0 = -(\mathbf{r}\times\nabla)\nabla' \int_0^a \frac{g(r',r'')}{4\pi} \sum_{n=0}^{\infty} \frac{(r'')^n}{r^{n+1}} P_n(\cos\gamma) dr''$$

$$= \nabla \times \mathbf{r}\nabla' \int_0^a \frac{g(r',r'')}{4\pi D(\mathbf{r}-\mathbf{r}'')} dr''. \tag{54}$$

Here, the sum in (50) has been extended from $n = 1$ to $n = 0$ because $\mathbf{r} \times \nabla P_0(\cos\gamma) = 0$.

Writing the first-order magnetic field as

$$\mathbf{H}_1 = j\omega[\nabla \times \overline{\overline{G}}^r]_0 \cdot \mathbf{p}_0 = \frac{1}{\mu_o}\nabla \times \mathbf{A}_1, \tag{55}$$

the vector potential can be defined as

$$\mathbf{A}_1(\mathbf{r}) = j\omega\mu_o \mathbf{p}_0 \cdot \nabla' \int_0^a \frac{\mathbf{r} g(r',r'')}{4\pi D(\mathbf{r}-\mathbf{r}'')} dr''. \tag{56}$$

Because

$$\frac{\mathbf{r}}{D(\mathbf{r}-\mathbf{r}'')} = \nabla D(\mathbf{r}-\mathbf{r}'') + \frac{\mathbf{r}''}{D(\mathbf{r}-\mathbf{r}'')}, \tag{57}$$

and since the gradient term vanishes in the curl expression of the magnetic field, it can be omitted and we can further write

$$\mathbf{A}_1(\mathbf{r}) = j\omega\mu_o \mathbf{p}_0 \cdot \nabla' \int_0^a \frac{\mathbf{r}'' g(r',r'')}{4\pi D(\mathbf{r}-\mathbf{r}'')} dr''$$

$$= j\omega\mu_o \int_0^a \frac{\mathbf{r}''(\mathbf{p}_0 \cdot \nabla' g)}{4\pi D} dr'' + j\omega\mu_o \int_0^a \frac{(\mathbf{p}_0 \cdot \nabla' \mathbf{u}_r')\mathbf{r}'' g}{4\pi D} dr''$$

$$+ j\omega\mu_o \int_0^a \frac{\mathbf{r}'' g}{4\pi} \mathbf{p}_0 \cdot \nabla' \frac{1}{D} dr''. \tag{58}$$

This expression can be compared with that for the vector potential due to a line current along the $z$ axis, $\mathrm{I}(z'')$:

$$\mathrm{A}(\mathrm{r}) = \mu_o \int_0^a \frac{\mathrm{I}(z'')}{4\pi D(\mathrm{r} - \mathrm{u}_z z'')}\, dz'', \tag{59}$$

and due to a bifilar line current defined by the current density function

$$\mathrm{J}(\mathrm{r}'') = \mathrm{I}(z'')\mathbf{L} \cdot \nabla'' \delta(\boldsymbol{\rho}''). \tag{60}$$

$\boldsymbol{\rho}$ is the transverse ($xy$ plane) position vector, $\mathbf{L}$ the infinitesimal distance vector between two infinite currents, $+\mathbf{I}$ flowing along the $z$ axis and $-\mathbf{I}$ parallel to it through the point $\boldsymbol{\rho}'' = \mathbf{L}$. $\mathbf{IL}$ is the finite moment dyadic of the bifilar current. The corresponding vector potential can be written as

$$\mathrm{A}(\mathrm{r}) = \mu_o \int_0^a \frac{\mathrm{J}(\mathrm{r}'')}{4\pi D(\mathrm{r} - \mathrm{u}_z z'')} dV'' = -\mu_o \int_0^a \frac{\mathrm{I}(z'')\mathbf{L}}{4\pi} \cdot \nabla'' \frac{1}{D} dz''. \tag{61}$$

To find the image currents, let us again separate the two cases:

**Radial dipole**

In this case, the operator $\mathrm{p}_0 \cdot \nabla'$ only operates on the function $g(r', r'')$, whence the last two integrals in (58) vanish. Comparing with (59), the vector potential can be expressed in terms of an image line current flowing along the $\mathrm{r}'$ direction:

$$\begin{aligned}
\mathrm{I}_{i1}(r'') &= j\omega \mathrm{p}_0 \cdot \nabla' g(r', r'') \mathrm{u}_r' r'' = j\omega \mathrm{p}_0 \frac{d}{dr'} [r'' g(r', r'')] \\
&= j\omega \mathrm{p}_0 \frac{d}{dr'} \left[ \frac{\epsilon_r - 1}{\epsilon_r + 1} \frac{r''}{a} \left( \frac{r' r''}{a^2} \right)^{-\frac{\epsilon_r}{\epsilon_r + 1}} U\left( \frac{a^2}{r'} - r'' \right) \right] \\
&= -j\omega \mathrm{p}_0 \left[ \frac{\epsilon_r - 1}{\epsilon_r + 1} \left( \frac{a}{r'} \right)^3 \delta(r'' - \frac{a^2}{r'}) \right. \\
&\quad \left. + \frac{\epsilon_r(\epsilon_r - 1)}{(\epsilon_r + 1)^2} \frac{a}{(r')^2} \left( \frac{r' r''}{a^2} \right)^{\frac{1}{\epsilon_r + 1}} U\left( \frac{a^2}{r'} - r'' \right) \right].
\end{aligned} \tag{62}$$

It is seen that the image consists of a line current between the origin and Kelvin point $r = a^2/r'$ and a dipole at the Kelvin point. Figure 2 depicts the continuous part of the normalized image current function for different distances of the dipole and $\epsilon_r = 5$.

The divergence of the image current

$$\nabla'' \cdot \mathrm{I}_{i1} = j\omega p_0 \frac{\partial^2}{\partial r' \partial r''} \left[ \frac{\epsilon_r - 1}{\epsilon_r + 1} \frac{r''}{a} \left( \frac{r' r''}{a^2} \right)^{-\frac{\epsilon_r}{\epsilon_r + 1}} U\left( \frac{a^2}{r'} - r'' \right) \right], \tag{63}$$

when compared with (44), shows that the condition of continuity

$$\nabla'' \cdot \mathbf{I}_{i1} = -j\omega\varrho_{i0} \tag{64}$$

is satisfied. This can be pictured so that the charges of the static image make up the current of the first-order image through sinusoidal motion along the image line.

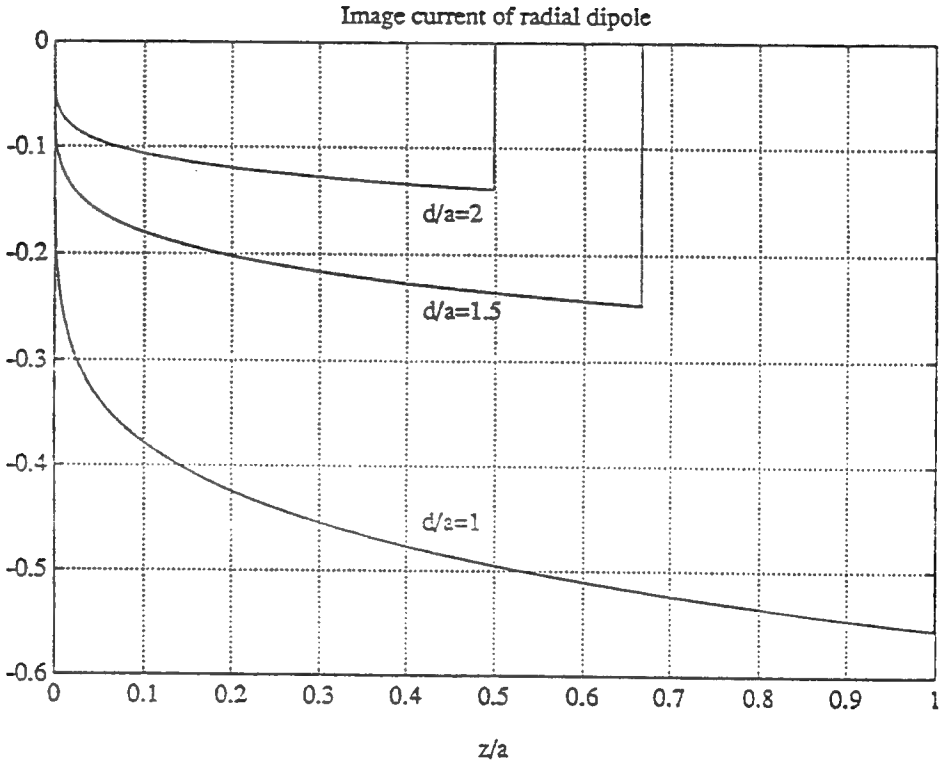



Figure 2. Normalized functions characterizing the image current line corresponding to a radial dipole at different distances $d$ for $\epsilon_r = 5$. The delta function at the Kelvin point is not shown.

Transverse dipole

In the case of $\mathbf{p}_0$ perpendicular to $\mathbf{r}'$ we have $\mathbf{p}_0 \cdot \nabla' g(r', r'') = 0$, and the last two integrals in (58) form the expression for the vector potential:

$$\begin{aligned}\mathbf{A}_1(\mathbf{r}) =& j\omega\mu_o \int_0^a \frac{(\mathbf{p}_0 \cdot \nabla' \mathbf{u}_\tau')r'' g}{4\pi D}\, dr'' + j\omega\mu_o \int_0^a \frac{\mathbf{u}_\tau' r'' g}{4\pi}\, \mathbf{p}_0 \cdot \nabla' \frac{1}{D}\, dr'' \\ =& \mu_o \int_0^a \frac{j\omega \mathbf{p}_0 r'' g}{4\pi r' D}\, dr'' + \mu_o \int_0^a \frac{j\omega \mathbf{u}_\tau' (r'')^2 g \mathbf{p}_0}{4\pi r'} \cdot \nabla'' \frac{1}{D}\, dr''. \end{aligned} \tag{65}$$

For the last step we have invoked the property

$$\mathbf{p}_0 \cdot \nabla' \frac{1}{D(\mathbf{r} - \mathbf{u}_\tau' r'')} = \frac{r''}{r'} \mathbf{p}_0 \cdot \nabla'' \frac{1}{D(\mathbf{r} - \mathbf{r}'')}. \tag{66}$$

Comparing (65) with (59) and (61) shows us that the image can be written as a combination of a transverse current strip plus a longitudinal bifilar line current. The first integral of (65) corresponds to a transverse current which has the form

$$\mathbf{I}_{i1}(r'') = j\omega \mathbf{p}_0 \frac{r''}{r'} g(r', r''). \tag{67}$$

The second term corresponds to a bifilar current along the direction $\mathbf{r}'$ with the moment dyadic

$$[\mathbf{II}]_{i1} = -j\omega \mathbf{u}_\tau' \frac{(r'')^2}{r'} \mathbf{p}_0 g(r', r''). \tag{68}$$

Thus, a transverse current element corresponds to an image which has both axial and transverse components. Combining these leads finally to the image current density

$$\begin{aligned}\mathbf{J}_{i1}(\mathbf{r}'') =& j\omega \mathbf{p}_0 \frac{r''}{r'} g(r', r'') \delta(\boldsymbol{\rho}'') - j\omega \mathbf{u}_\tau' \frac{(r'')^2}{r'} g(r', r'') \mathbf{p}_0 \cdot \nabla'' \delta(\boldsymbol{\rho}'') \\ =& j\omega \frac{r''}{r'} g(r', r'') \nabla'' \times \left[\mathbf{p}_0 \times \mathbf{u}_\tau' r'' \delta(\boldsymbol{\rho}'')\right]. \end{aligned} \tag{69}$$

Let us again check the continuity condition. The divergence of the image current density (69) can be written as

$$\nabla'' \cdot \mathbf{J}_{i1} = j\omega \frac{r''}{r'} f(r', r'') \nabla'' \cdot [\mathbf{p}_0 \delta(\boldsymbol{\rho}'')] = j\omega \nabla'' \cdot \mathbf{p}_{i0} = -j\omega \varrho_{i0}. \tag{70}$$

The last term refers to the image charge density corresponding to the image dipole density given in the static case (45). It is seen that also in the transverse dipole case the first-order image current and the zeroth-order charge satisfy the continuity condition, which means that the first-order image current can be pictured as being the zeroth-order image charge in periodic motion. Since the current cannot

be determined from the charge, because it is not uniquely determined from its divergence, the present analysis was needed to find the result.

The image expression for the transverse dipole (69) can also be written in the form

$$\mathbf{J}_{i1}(\mathbf{r}'') = \nabla'' \times \left[ j\omega \frac{(r'')^2}{r'} g(r', r'') \mathbf{p}_0 \times \mathbf{u}_r' \delta(\boldsymbol{\rho}'') \right] + j\omega \frac{r''}{r'} f(r', r'') \mathbf{p}_0 \delta(\boldsymbol{\rho}''), \quad (71)$$

of which the first term can be expressed in terms of an equivalent magnetic current [11]:

$$\mathbf{J} = \frac{\nabla \times \mathbf{J}_{me}}{j\omega\mu_o}. \quad (72)$$

Thus, as an alternative to the expression (69), the transverse dipole $\mathbf{p}_0$ can be seen to give rise to transverse electric and magnetic image strip currents of the form

$$\mathbf{I}_{i1} = j\omega \frac{r''}{r'} f(r', r'') \mathbf{p}_0, \quad (73)$$

$$\mathbf{I}_{mi2} = -\frac{k_o^2}{\epsilon_o} \frac{(r'')^2}{r'} g(r', r'') \mathbf{p}_0 \times \mathbf{u}_r'. \quad (74)$$

These two strip currents are seen to be at right angles to one another. Their normalized functional dependence is depicted in Fig. 3 for certain relative permittivities of the sphere.

## 4. SUMMARY OF RESULTS

Let us summarize the first-order low-frequency image results in cartesian coordinate form which is more easily applicable in further analysis. The indices showing the order of the image are omitted in the sequel.

For a slowly oscillating current dipole on $z$ axis, defined by

$$\mathbf{J}(\mathbf{r}) = \mathbf{u} IL \delta(z - d) \delta(\boldsymbol{\rho}), \quad (75)$$

the image corresponding to the axial dipole, with $\mathbf{u} = \mathbf{u}_z$, is

$$\mathbf{J}_i(\mathbf{r}) = -\mathbf{u}_z IL \left[ \frac{\epsilon_r - 1}{\epsilon_r + 1} \left( \frac{a}{d} \right)^3 \delta\left( z - \frac{a^2}{d} \right) + \frac{\epsilon_r(\epsilon_r - 1)}{(\epsilon_r + 1)^2} \frac{a}{d^2} \left( \frac{zd}{a^2} \right)^{\frac{1}{\epsilon_r + 1}} U\left( \frac{a^2}{d} - z \right) \right] \delta(\boldsymbol{\rho}) \quad (76)$$

while the transverse dipole, with $\mathbf{u} \cdot \mathbf{u}_z = 0$, corresponds to the image current density

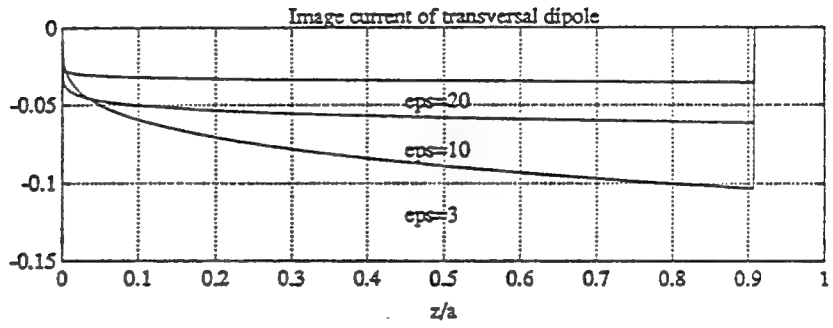
$$\mathbf{J}_i(\mathbf{r}) = IL \frac{\epsilon_r - 1}{\epsilon_r + 1} \frac{a}{d^2} \left( \frac{zd}{a^2} \right)^{\frac{1}{\epsilon_r + 1}} U\left( \frac{a^2}{d} - z \right) [\mathbf{u} \delta(\boldsymbol{\rho}) - z \mathbf{u}_z \mathbf{u} \cdot \nabla \delta(\boldsymbol{\rho})], \quad (77)$$

or to the combination of electric and magnetic image current densities

$$\mathbf{J}_i(\mathbf{r}) = \mathbf{u}IL\left[\frac{\epsilon_r - 1}{\epsilon_r + 1}\left(\frac{a}{d}\right)^3 \delta\left(z - \frac{a^2}{d}\right) - \frac{\epsilon_r - 1}{(\epsilon_r + 1)^2}\frac{a}{d^2}\left(\frac{zd}{a^2}\right)^{\frac{1}{\epsilon_r+1}} U\left(\frac{a^2}{d} - z\right)\right]\delta(\rho), \tag{78}$$

$$\mathbf{J}_{mi}(\mathbf{r}) = (\mathbf{u} \times \mathbf{u}_z)j\omega\mu_o IL\frac{\epsilon_r - 1}{\epsilon_r + 1}\left(\frac{a}{d}\right)^3\left(\frac{zd}{a^2}\right)^{\frac{\epsilon_r+2}{\epsilon_r+1}} U(\frac{a^2}{d} - z)\delta(\rho). \tag{79}$$

(a)




(b)

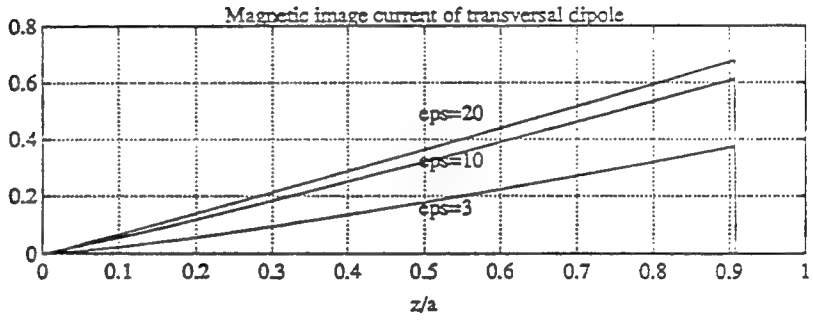



Figure 3 Normalized electric (a) and magnetic (b) components of the image current corresponding to a transverse dipole at the distance $d = 1.1a$ for different values of the relative permittivity of the sphere, $\epsilon_r$. The currents flow transverse to the $z$ axis and each other. The delta function at the Kelvin point is not shown.

## 5. SPECIAL CASES

Let us finally study some limit cases to check the image current expressions (76)–(79) for the dielectric sphere.

1. For $\epsilon_r \to 1$ the sphere vanishes. From (76)–(79) we see that all image currents vanish because of the factor $\epsilon_r - 1$.

2. For $\epsilon_r \gg 1$ the dielectric sphere becomes somewhat similar to an ideally conducting sphere for the electric field. In the present case, for a radial dipole, the image current (76) has the form

$$\mathbf{J}_i = -j\omega \mathbf{u} IL \left[ \left(\frac{a}{d}\right)^3 \delta\left(z - \frac{a^2}{d}\right) + \frac{a}{d^2} U\left(\frac{a^2}{d} - z\right) \right] \delta(\rho), \tag{80}$$

i.e., it is composed of a point dipole plus a constant line current. This simple result does not seem to be known. For a transverse dipole, the image current density (77) is

$$\mathbf{J}_i = j\omega IL \frac{a}{d^2} \nabla \times \left[\mathbf{u} \times \mathbf{u}_z z \delta(\boldsymbol{\rho}'')\right] U\left(\frac{a^2}{d} - z\right), \tag{81}$$

which consists of a constant transverse current plus a bifilar axial line current, whose amplitude is proportional to $z$. From (78), (79) we have another representation in terms of an electric and a magnetic current

$$\mathbf{J}_i = j\omega \mathbf{u} IL \left(\frac{a}{d}\right)^3 \delta\left(z - \frac{a^2}{d}\right) \delta(\rho), \tag{82}$$

$$\mathbf{J}_{mi} = -(\mathbf{u} \times \mathbf{u}_z) \frac{k_o^2}{\epsilon_o} IL \frac{az}{d^2} U\left(\frac{a^2}{d} - z\right) \delta(\rho). \tag{83}$$

This means that the other form for the image of the transverse dipole consists of a transverse dipole plus a transverse magnetic line current with linear amplitude dependence. Note that application of these results requires a small enough frequency because in deriving the first-order theory it was assumed that $\sqrt{\epsilon_r}\, k_o r$ be small.

3. For $d \gg a$ we have a dielectric sphere in homogeneous incident electric field. Because the distance to the Kelvin point $a^2/d$ becomes small, the image is concentrated at the center. Moments of the image give multipole terms for the image point source.

For the radial dipole we can write from (76) for the zeroth moment of the image current, i.e., the moment of the image dipole at the center of the sphere, the expression

$$\int_V \mathbf{J}_i(\mathbf{r}) dV = -2\mathbf{u}_z IL \frac{\epsilon_r - 1}{\epsilon_r + 2} \left(\frac{a}{d}\right)^3. \tag{84}$$

From (77) we find the image moment corresponding to the transverse dipole with direction $\mathbf{u}$:

$$\int_V \mathbf{J}_i(\mathbf{r})dV = \mathbf{u}IL\frac{\epsilon_r - 1}{\epsilon_r + 2}\left(\frac{a}{d}\right)^3. \tag{85}$$

Because the lowest-order terms for the electric field incident from a radial and a transverse dipole with the respective moments $\mathbf{u}_z IL$ and $\mathbf{u}IL$ at the distance $d$ are, respectively,

$$\mathbf{E}^{inc} = \frac{2\mathbf{u}_z IL}{4\pi\epsilon_o d^3}, \qquad \mathbf{E}^{inc} = -\frac{\mathbf{u}IL}{4\pi\epsilon_o d^3}, \tag{86}$$

these inserted in the respective expressions (84) and (85) give the same dipole moment expression

$$\mathbf{I}_i L = -4\pi\epsilon_o a^3 \frac{\epsilon_r - 1}{\epsilon_r + 2}\mathbf{E}^{inc} \tag{87}$$

responsible for the lowest-order scattered dipolar field. This coincides with the well-known result obtained elsewhere through the Stevenson analysis [14].

There is a magnetic moment corresponding to the image of the transverse dipole, which is obtained by integrating (79) over the sphere:

$$\int_V \mathbf{J}_{mi}(\mathbf{r})dV = (\mathbf{u}\times\mathbf{u}_z)j\omega\mu_o IL\frac{\epsilon_r - 1}{2\epsilon_r + 3}\frac{a^5}{d^4}. \tag{88}$$

This is of smaller order than the electric moments because of the $1/d^4$ dependence on the distance. Simple expressions for higher order multipole moments can also be readily obtained from the image current expressions (76) and (77).

4. For $a \to \infty$ the spherical interface becomes planar and we can compare the result with that given by the Exact Image Theory [15] in the low-frequency case.

Taking, for example, the transverse magnetic (TM) image current due to a vertical electric dipole above a dielectric half space from [16], we can write for the exact image current

$$I_i^{TM}(\zeta) = -\mathbf{u}_z IL f^{TM}(\zeta) \tag{89}$$

with

$$f^{TM}(\zeta) = -IL\left[\frac{8\epsilon_r}{\epsilon_r^2 - 1}\sum_{n=1}^{\infty} n\left(\frac{\epsilon_r - 1}{\epsilon_r + 1}\right)^n \frac{J_{2n}(jB\zeta)}{\zeta} - \frac{\epsilon_r - 1}{\epsilon_r + 1}\delta(\zeta)\right], \tag{90}$$

$$B = k_o\sqrt{\epsilon_r - 1}, \qquad k_o = \omega\sqrt{\mu_o\epsilon_o}. \tag{91}$$

The variable $\zeta$ measures distance from the mirror image point $\mathbf{r} = -\mathbf{u}_z h$ downwards. Because $f^{TM}(\zeta)$ is an even function of $B$ and, hence, of $\omega$, the low-frequency approximation $\beta \to 0$ starts with

$$f^{TM}(\zeta) \to IL\frac{\epsilon_r - 1}{\epsilon_r + 1}\delta(\zeta) + IL\frac{\epsilon_r(\epsilon_r - 1)}{(\epsilon_r + 1)^2}k_o^2\zeta U(\zeta) + \ldots \tag{92}$$

Thus, in the first order approximation, only the delta term remains and the image current is

$$I_i^{TM}(\zeta) \approx IL \frac{\epsilon_r - 1}{\epsilon_r + 1} \delta(\zeta). \tag{93}$$

In the present sphere problem, taking the limit $a \to \infty$ and denoting $d = h+a$, $\zeta = a-(z+h)$, where $h$ is the height of the dipole from the interface, substituting $z = a - (h + \zeta)$, and $a^2/d \to a - h$ in (76), (77) leaves us with the asymptotic image expression

$$I_i \to IL \left[ \frac{\epsilon_r - 1}{\epsilon_r + 1} \delta(\zeta) + \frac{\epsilon_r(\epsilon_r - 1)}{(\epsilon_r + 1)^2} \frac{1}{a} U(\zeta) \right], \tag{94}$$

of which the second terms vanishes and the first one coincides with the exact image result above.

## ACKNOWLEDGMENTS

The present study was mainly made while the first author was on sabbatical leave at the University of Delaware on a research fellowship from the Academy of Finland in April 1992. Support from the Academy of Finland has also helped in finishing the paper.

The Editor thanks two anonymous Reviewers for reviewing the paper.

## REFERENCES

1. W. Thomson (Kelvin), "Extracts of two letters to mr. Liouville" (in French), *Journal de Mathématiques Pures et Appliqués,* Vol. 10, 1845, 364; Vol. 12, 1847, 256.
2. Dassios, G., R. E. Kleinman, "On Kelvin inversion and low-frequency scattering", *SIAM Rev.,* Vol. 31, No. 4, 565–585, 1989.
3. Sezginer, A., W. C. Chew, "Image of a static current loop over a superconducting sphere", *IEEE Trans. Magnetics,* Vol. 26, No. 3, 1137–1138, May 1990.
4. Lindell, I. V., "Electrostatic image theory for the dielectric sphere," *Radio Sci.,* Vol. 27, No. 1, 1–8, Jan.-Feb. 1992.
5. Lindell, I. V., E. A. Lehtola, "Magnetostatic image theory for the permeable sphere," *IEEE Trans. Magnetics,* Vol. 28, No. 4, 1930–1934, July 1992.
6. Lindell, I. V., M. E. Ermutlu, A. H. Sihvola, "Electrostatic image theory for the layered dielectric sphere," *IEE Proc.,* Vol. H-139, No. 2, 186–192, April 1992.
7. Lindell, I. V., J. C-E. Sten, K. I. Nikoskinen, "Electrostatic image solution for the interaction of two dielectric spheres," *Radio Sci.,* Vol. 28, No. 3, 319–329, May-June 1993.
8. Lindell, I. V., "Quasi-static image theory for the bi-isotropic sphere," *IEEE Trans. Antennas Propag.,* Vol. 40, No. 2, 228–233, Feb. 1992.
9. Sten, J. C-E., I. V. Lindell, "Electrostatic image theory for the dielectric sphere with an internal source", *Microwave Opt. Tech. Lett.,* Vol. 5, No. 11, 597–602, October 1992.
10. Neumann, C., *Hydrodynamische Untersuchungen,* Leipzig: Teubner, 1883.
11. Mayes, P. E., "The equivalence of electric and magnetic sources", *IRE Trans. Antennas Propagat.,* Vol. 6, No. 4, 295–296, July 1958.

12. Stevenson, A. F., "Solution of electromagnetic scattering problems as power series in the ratio (dimension of scatterer)/(wavelength)," *Jour. Appl. Phys.*, Vol. 24, 1134–1142, 1953.
13. Jones, D. S., *The Theory of Electromagnetism*, Pergamon, 1964, pp.495-500.
14. Bladel, J. van, *Electromagnetic Fields*, New York: Mc Graw-Hill, 1964, pp.282-284.
15. Lindell, I. V., E. Alanen, "Exact image theory for the Sommerfeld half-space problem. Part II: Vertical electric dipole," *IEEE Trans. Antennas Propagat.*, Vol. 32, No. 8, 841–847, August 1984.
16. Lindell, I. V., *Methods for Electromagnetic Field Analysis*, Oxford: Oxford University Press, 1992.

**Ismo V. Lindell** born in Viipuri, Finland (1939) and received the Dr. Tech. (1971) in Helsinki University of Technology, Finland, where he is presently Professor of Electromagnetic Theory and Director of the Electromagnetics Laboratory. He took sabbaticals at the University of Illinois (1972-73) and M.I.T. (1986-87). Presently, he is on the URSI National Committee and he is the Commission B Chairman of Finland and AdCom Member of the IEEE Antennas and Propagation Society. He is a Fellow of the IEEE, one of the Editors of *Journal of Electromagnetic Waves and Applications*, and an Associate Editor of *Radio Science*. Dr. Lindell has authored and co-authored *Methods for Electromagnetic Field Analysis* (1992), and *Electromagnetic Waves in Chiral and Bi-Isotropic Media* (to appear in 1994). He is the recipient of the 1987 S.A. Schelkunoff award of the IEEE Antennas and Propagation Society.

**Johan Sten** was born in Helsinki, Finland, in 1967. He received the Dipl.Eng. degree in Electrical Engineering from the Helsinki University of Technology, Espoo, Finland in 1992. Currently he is with the Electromagnetics Laboratory of the H.U.T. His interests are in analytical methods, especially image theories, in electromagnetics.

**Ralph Kleinman** was born in 1929, received a B.A in Mathematics from New York University in 1950, an M.A. in Mathematics from the University of Michigan in 1951, and a Ph.D. in Mathematics from the Technische Hogeschool, Delft, the Netherlands, in 1961. In 1951, he joined the group that later became the Radiation Laboratory at the University of Michigan. He remained at the Radiation Laboratory until 1968, when he joined the Department of Mathematical Sciences at the University of Delaware, where he presently is a Professor. His main interests have been concerned with the mathematical problems associated with propagation and scattering of acoustic and electromagnetic waves including radar cross section analysis, integral representations of solutions of the Eelmholtz and Maxwell equations, and low frequency perturbation techniques, and has guided the work of doctoral students in these areas.